

```

1 AAGCGATAGC TGAGTGC GGC TGCTGAT TGTGTTCTAG GGGACGGAGT
51 AGGGGAAGAC GTTGTCTCTC CCGGAACAGC CTATCTCATT CCTTCTTTTC
101 GATTACCCGT GGC GCGGAGA GTCAGGCGG CGGCTGCGGC AGCAAGGGCG
151 GCGGTGGCGG CCGCGGCAGC TGCAGTGACA TGTCCAGCAT GAATCCCGAA
201 TATGATTATT TATTCAAGTT ACTTCTGATT GCGCACTCAG GGGTTGGAAA
251 GTCTTGCCTT CTTCTTAGGT TTGCAGATGA TACATATACA GAAAGCTACA
301 TCAGCACAAAT TGGTGTGGAT TTCAAAATAA GAACTATAGA GTTAGACGGG
351 AAAACAATCA AGCTTCAAAT AGAGTCCTTC AATAATGTTA AACAGTGGCT
401 GCAGGAAATA GATCGTTATG CCAAGTAAAA TGTCAACAAA TTGTGGTAG
451 GGAACAAATG TGATCTGACC ACAAAGAAAG TAGTAGACTA CACAACAGCG
501 AAGGAATTTG CTGATTCCCT TGAATTCCTG TTTTGGAAA CAGTGCTAA
551 GAATGCAACG AATGTAGAAC AGTCTTTCAT GACGATGGCA GCTGAGATTA
601 AAAAGCGAAT GGGTCCCGGA GCAACAGCTG GTGGTGCTGA GAAGTCCAAT
651 GTTAAAAATC AGAGCACTCC AGTCAAGCAG TCAGGTGGAG GTTGCTGCTA
701 AAATTTGCCT CCATCCTTTT CTCACAGCAA TGAATTTGCA ATCTGAACCC
751 AAGTGAAAAA ACAAATTCG CTGAATTGTA CTGTATGTAG CTGCACTACA
801 ACAGATTCTT ACCGTCTCCA CAAAGGTCAG AGATTGTAAA TGGTCAATAC
851 TGACTTTTTC TTTATTCCTT TGAAGTCAAG CAGCTAAGT CATTTTCAGA
901 ACTGTTTTAA ACCTTTGTGT GCTGGTTTAT AAAATAATGT GTGTAATCCT
951 TGTGCTTTTC CTGATACCAG ACTGTTTCCC GTGGTTGGTT AGAATATATT
1001 TTGTTTTGAT GTTTATATTG GCATGTTTAG ATGTCAGGTT TAGTCTTCTG
1051 AAGATGAAGT TCAGCCATTT TGTATCAAAC AGCACAAGCA GTGTCTGTCA
1101 CTTTCCATGC ATAAAGTTA GTGAGATGTT ATATGTAAGA TCTGATTTCG
1151 TAGTCTTCC TTGTAGAGTT ATAAATGGAA AGATTACACT ATCTGATTAA
1201 TAGTTCTTTC ATACTCTGCA TATAATTTGT GGCTGCAGAA TATGTAAATT
1251 TGTGTCACAC TATGTAACAA AACAACTGAA GATATGTTTA ATAAATATTG
1301 TACTTATTGG AAGTAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
1351 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
1401 AAAAA (SEQ ID NO:1)

```

FEATURES:

```

5'UTR:      1-179
Start Codon: 180
Stop Codon:  699
3'UTR:      702

```

Homologous proteins:

Top 10 BLAST Hits

	Score	E
CRA 108000024647144 /altid=gi 12728868 /def=ref XP_002675.2 RA...	372	e-102
CRA 18000004923424 /altid=gi 4758988 /def=ref NP_004152.1 RAB1...	332	5e-90
CRA 18000004937406 /altid=gi 131787 /def=sp P05711 RB1A_RAT RAS...	328	1e-88
CRA 18000004952860 /altid=gi 131785 /def=sp P22125 RAB1_DISOM R...	320	3e-86
CRA 18000004995539 /altid=gi 103720 /def=pir D38625 GTP-bindin...	313	3e-84
CRA 18000004967528 /altid=gi 92339 /def=pir S06147 GTP-binding...	297	2e-79
CRA 18000004880958 /altid=gi 464524 /def=sp Q05974 RAB1_LYMST R...	282	9e-75
CRA 18000004908714 /altid=gi 466171 /def=sp P33723 YPT1_NEUCR G...	253	3e-66
CRA 18000005175724 /altid=gi 7497231 /def=pir T33781 hypotheti...	253	4e-66
CRA 335001098696672 /altid=gi 11558649 /def=emb CAC17833.1 (AJ...	251	2e-65

FIGURE 1, page 1 of 2

BLAST dbEST hits:

	Score	E
gi 12867866 /dataset=dbest /taxon=960...	654	0.0
gi 12097820 /dataset=dbest /taxon=96...	654	0.0
gi 12793758 /dataset=dbest /taxon=960...	624	e-177
gi 12338056 /dataset=dbest /taxon=96...	622	e-176
gi 11977068 /dataset=dbest /taxon=96...	609	e-172
gi 10339840 /dataset=dbest /taxon=960...	517	e-145
gi 10349761 /dataset=dbest /taxon=960...	436	e-120
gi 10997958 /dataset=dbest /taxon=96...	385	e-105
gi 10996533 /dataset=dbest /taxon=96...	381	e-103

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

From BLAST dbEST hits:

gi|12867866 Fetal brain
gi|12097820 Adrenal gland
gi|12793758 Brain neuroblastoma cell line
gi|12338056 Adrenal gland
gi|11977068 Skin melanotic melanoma
gi|10339840 Uterus leiomyosarcoma
gi|10349761 Skin melanotic melanoma
gi|10997958 Placenta
gi|10996533 Placenta

From tissue screening panels:

Whole brain

1 MSSMNPEYDY LFKLLIGDS GVGKSCLLLR FADDTYTESY ISTIGVDFKI
 51 RTIELDGKTI KLQIESFNNV KQWLQEIDRY ASENVNKLLV GNKCDLTTKK
 101 VVDYTTAKEF ADSLGIPFLE TSAKNATNVE QSFMTMAAEI KCRMGPATA
 151 GGAEKSNVKI QSTPVKQSGG GCC (SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
 N-glycosylation site

125-128 NATN

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
 Protein kinase C phosphorylation site

Number of matches: 5

1	59-61	TIK
2	97-99	TTK
3	98-100	TKK
4	106-108	TAK
5	122-124	SAK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
 Casein kinase II phosphorylation site

Number of matches: 3

1	35-38	TYTE
2	106-109	TAKE
3	127-130	TNVE

[4] PDOC00007 PS00007 TYR_PHOSPHO_SITE
 Tyrosine kinase phosphorylation site

30-36 RFADDTY

[5] PDOC00008 PS00008 MYRISTYL
 N-myristoylation site

Number of matches: 3

1	21-26	GVGKSC
2	147-152	GATAGG
3	152-157	GAEKSN

[6] PDOC00017 PS00017 ATP_GTP_A
 ATP/GTP-binding site motif A (P-loop)

18-25 GDSGVGKS

[7] PDOC00579 PS00675 SIGMA54_INTERACT_1
 Sigma-54 interaction domain ATP-binding region A signature

14-27 LLLIGDSGVGKSCL

BLAST Alignment to Top Hit:

>CRA|108000024647144 /altid=gi|12728868 /def=ref|XP_002675.2| RAB1,
member RAS oncogene family [Homo sapiens] /org=Homo
sapiens /taxon=9606 /dataset=nraa /length=222
Length = 222

Score = 372 bits (944), Expect = e-102
Identities = 190/222 (85%), Positives = 190/222 (85%), Gaps = 32/222 (14%)
Frame = +3

Query: 129 GGCGSKGGGGGGSCSDMSSMNPEYDYLFKLLIGDSGVGKSCLLRFADDTYTESYIST 308
GGCGSKGGGGGGSCSDMSSMNPEYDYLFKLLIGDSGVGKSCLLRFADDTYTESYIST
Sbjct: 1 GGCGSKGGGGGGSCSDMSSMNPEYDYLFKLLIGDSGVGKSCLLRFADDTYTESYIST 60

Query: 309 IGVDFKIRTIELDGKTIKLQI-----ESFNNVK 392
IGVDFKIRTIELDGKTIKLQI ESFNNVK
Sbjct: 61 IGVDFKIRTIELDGKTIKLQIWDTAGQERFRTITSSYYRGAGHIIVVYDVTQESFNNVK 120

Query: 393 QWLQEIDRYASENVNKLVLGNKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQ 572
QWLQEIDRYASENVNKLVLGNKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQ
Sbjct: 121 QWLQEIDRYASENVNKLVLGNKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQ 180

Query: 573 SFMTMAAEIKKRMGPGATAGGAEKSNVKIQSTPVKQSGGGCC 698
SFMTMAAEIKKRMGPGATAGGAEKSNVKIQSTPVKQSGGGCC
Sbjct: 181 SFMTMAAEIKKRMGPGATAGGAEKSNVKIQSTPVKQSGGGCC 222 (SEQ ID NO:4)

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00071	Ras family	256.4	7.7e-75	2
CE00060	CE00060 rab_ras_like	170.0	3.9e-47	2
PF00634	BRCA2 repeat.	9.9	0.39	1
PF00056	lactate/malate dehydrogenase	3.9	3.4	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00056	1/1	13	29 ..	1	18 [.	3.9	3.4
CE00060	1/2	8	64 ..	20	77 ..	86.8	8.9e-23
PF00071	1/2	13	64 ..	1	52 [.	111.9	4.8e-32
PF00634	1/1	57	79 ..	13	35 .]	9.9	0.39
CE00060	2/2	65	140 ..	110	188 ..	81.2	2.9e-21
PF00071	2/2	65	173 .]	85	198 .]	142.4	4.5e-41

1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378</
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

[REDACTED]

```

3151 TCACCACAGC AGATGTGACA CACCTGGCAC CTTTCCTGGG AACTGGTGT
3201 CACTTCCCTT GGGTAGAGTT TGTTGGGCTC TCCTCAATGG CCCTTTAAAA
3251 ATTTCCTCTA CAGTTTACAT GCATGTAAAG TAATGAATAA TTGGAAGAGA
3301 CCGAATTGGT ATTCCTTTTC AGTGTCAAAG GCCTTTGAGG GATGGGGGAA
3351 AATCAGTATT TGTTGTAAAA GTTGAGTTTA TTTGCTGGT TGGTCAATTA
3401 CTGCTAGACA TTTTCCCTTA AAAGGTCCAC CCACCAGTT AGCTGACTGT
3451 CATATGTGTG TCACATGGCT CTTGCAAAAT GCTTACAAGT TTTGTAATAG
3501 TGTGGCTTGA AGCTGAAATC TTTTGCAC TAACAGAAACC GTAGTATTTT
3551 ATTAGAATTT CATGCTTTAG AAGTTGAGGG TAGTGTCTT GTAGTGACAT
3601 TTGCTGTGTT GACAGTTTAA AAAAAATTTT TTTTCAAGG CTCCAAGGAC
3651 AAAGTTGGTT TTGCACAGTT GAACGGAGGT GAAGTTGAGG TTCTTAATTT
3701 AGTAGTTTTC TTGGTAACAA TAAAGAACAT GGATTTACTG CTTTATCGAG
3751 GTTTATAGAC CTCTACTGTT CAGGAAATTT TCTGAATTG CTATATATAT
3801 GTTTATTAGT GTAAATAAAT CTTCAAGATT AGTTGAGAAC TTTGACAAGT
3851 TACTCAGCCT CTGAATTTTT TTTCCCTTTT GTAAATAGG ATAATTGGAG
3901 TCATTATTC TGTCAGGGTA GTGGTGAAAT TCAAATGTAT ATAAAAGAAT
3951 TTGAAAAACT GTGTGAGCAT TCTTCAGGTG GTATGCATCA TTTTCATGAA
4001 AGGCATTCTA TTAGTACCAG GATTTAGGAA TATAATCCTT GCGCTTAAGA
4051 AGTTTAGATA TAGGCCAGGC GCGGTGGCTC ACCTCAGTAA TCCCAGCACT
4101 TTGGGAGGCC GAGGCGGGCG GATCCCGAGG TCAGGAGATC GAGACCATCC
4151 TCGGTAACAC GGTGAAACCC CGTCTCTACT AAAAATGCAA AAAAAATAGC
4201 CGGGCTGGTT GGTGGGCACC TGTAAGTCCA GCTACTCGAG AGGCTGAGGC
4251 AGGAGAATGG CGTGATCCCG GGAGGTGGAG CTTGCAGTGA ACCAAGATCT
4301 GGCCACTGCA CTCCAGCCTG GACGACAGAG CAAGACTCCG TCTCAAAAAA
4351 AAAATTATTT ATTGTTTGA GACGGAGTTT CAATCTTGT GCCCAGGCTG
4401 GAGTGCAATG GCGCAAATCT CCTCTCACCG CCACCTCCGC CTCTGGGTT
4451 CAAGTGATTC TCCTGCCTCA GATTCCCGAG AAGTTGGGAT TACAGGCATG
4501 TGCCACCACT CCCGGCTAAT TTTGTATTTT TGGTAGAGAC GGGGTTTCTC
4551 CATGTTGGTC AGGCTGGTCT CAAACTCCCG AAGTGATCCG CCCGCTCAG
4601 CTTCCCAAAG TGTTGGGATT ACAGGCGTGA GCCACCGCGC CCGGCAGAAA
4651 TAGATTTTAT ACATGTCAAA TACCAGTAGA TATAGCAAT TCCAGATGTG
4701 TGGCATGGAT GAGAGCAACA AGATTTCAGG GGGATGGTGG GTTGTGGTTG
4751 GCTATCTGGG TTTTGGGAAGA CTTTATAGAA GAGAGACCTG AAAGGGATTT
4801 ATCAGCAATT AGATTTGGAG GAACAGAGGG AGTGACTAGG AATTTTCAAG
4851 GGGGAGAAGA AGGAGGAATG GCTCATAAAT GACAAGGACA GTAATAAGTA
4901 AATACGGTGT CAAATCATCC TTTCTTTTGA AGACTAATGA CCTCAAAGGG
4951 ATCAAACCCA GAAACAGTTT TTATATTTTT TCTGGGATCA AATACATGGG
5001 TATCTGGCCT ACTATATTTG TATTCTAGAC TGTTTAGTAA AATAATACAG
5051 GAATTTGAGA AAACCTTTGC AAAAGTGTTA GTGAAAAATTA CTAGGGTGA
5101 GAGGAAGTGA GGGATATTTT ATTAGGGGAG GTCACAAGGG CAGTGAGCAA
5151 TCAGATTTT AGTAATCTGA CTTAAGCAGT TTCTTTTTGT TTTAATGAAG
5201 CTGTGTTATCT TTATAAAAGT AATTAGAGAA AATTTGGAAA ATAAAGGAAA
5251 GAAAGAAAAG TTCTTTAGTG TTTTATCACG CAAATACAAG CTCAATCGTT
5301 TTTAACATCT TGTTCCAAAC TCCAAAGTCT TGCTTTCTCT TCAATTAATA
5351 CTTTAATGGG TGGATGCTTT TCCTGCTTCC AGTATGTTAT CTTAATAACT
5401 AACAATGGTA TATTAGCTAA TGTTTACAAA TGTAATCCAG ATGTTCTTAA
5451 AGTTACTTTG GTTTATCATT ACCAATTTAT ATTGTTTCTT TTAGAAATTT
5501 ATAATCTTTG TTAATGGGTT CTGCTAAATT TGGTAGTGAA AATGGGATCT
5551 TGAGAAAAAA GATTCTGAAG CAACAGAATT TTTAGATTTA TATTGGTTTA
5601 CATAAGAGTT GGTAGCTGTA TTAATTTTTT TGTTTGTGTT GTTTTTTTTT
5651 TGAGACGGAA TCTTGCTCTG TCGCCAGGC CTTGGCCTCC CAAAGTGTTG
5701 GGATTACAGG CGTGAGCCAC TGTGCTGGC TGTTTGTGTT TTTTTTGT
5751 TTTGTTTTCT TTTCTTTTTC TTTTTCCTGA GATGGAGTCT CACTCTGTCA
5801 CCCAGGCTGG AGTGCACTGG CGCGATCTTG GCTCACTGCA ATCTCTGCCT
5851 CCTGGGTTCA AGCGATTTTC CTGCCTTGGT CTCCTGAGTA GCTGGGATTA
5901 CAGGCATTTG CCACCATAAC CAGCTAATTT TTGTATAGAG TACCCAGCCA
5951 TCTCTAATGT TGATCAGGCT GAAGCAGGTG GATCACCTAA GGTCAAGGAT
6001 TCAAGACCAG CCTGGCCAAT ATGGCAAAAC CCTATCTCTA CTAATACAGA
6051 AAATTATCTG GGTGTGTTGG CTGGCGCCTG TAATCCAGC TACTCGGAG
6101 GCTGAGGCAG GACAATCTCT TGAACCTCGG AGGTGGAGGT TGCAGTGAGC
6151 CGAGATCACA CCATTGCACT CCAGCTGGG CAACAGAGCA AGACTTGCTCT
6201 CAAAAAAGGC AAAAAAAGGC AAAAAAAGGC AATTGAAAGT GTAATCTGAA
6251 CAGTTAAAAA AGTAGATAGA AAGGTTTAAA GCTTTTTTTT GAGGATCTGA

```

FIGURE 3, page 2 of 21

6301 AGAAAAATGT GGATTTTTTT TGAGCTACGT TTTGAAGCAG GCAGTGATTA
6351 TTTCAGCACA TTAAGAAATG CTTAACATGG CCAGGCGCAG TGGCTCACGC
6401 CTGTAATTCT CAGCACTTTG GGAGGCCGAG GTGGGCGGAT CATTTGAGGT
6451 CATGACCAGC CTGGCCAACA TGATGAGACA CTGCCTCTAC TAAAAATACA
6501 AAAATTAGCT GGGTGTGGTG GTGCACGCCT GTAATTCCAG CTAAGTACGA
6551 ACCTGAGGCA GGAGAGTCAC TTGAACCTGG GAGGCGGAGG CTGCAGTGAG
6601 TCCAGATCAT GCCACTGCAC TCCAGCCTGA GGGACAGAGT GAGACTCCTC
6651 AAAAAAAGAA AAAAAAAGAA AAAGAAATAC TTAACATTAT TCTCGTGATT
6701 ATTCTCATAA CATTTTTTCAT AATCCACTGG CTTCAGTGG ATTTTTTTAG
6751 TGCAAGAAA ATAATTTTGA TTGGTTCATC TTTAAGGAAT GTGTTAAGAA
6801 TAAAGCATGT CTACCTGTCT TCAGTATACC AGCTAACTAT AGTAGGAAGA
6851 AATATAGTAG TCTACTTAGA TCAACTATAA TTCTTTAATG CAGAAAAAGT
6901 TTAAAGTATT TACCTTATTT TTAGCCCCCA TCCCCTTAAG TATATCATGG
6951 TCCAGAAATC TCTGAAAATG TTATCAGTCT TTCAAGCTTT GCTCTTCTTT
7001 CATGTTATAC TCAAGAAACA TTTGACCTTT TTTTTTTTTT TTTTGCTTGC
7051 ATTGTGTTTC AAATAATTTT TAACAAAACCT TAAGTGTGTTG AAAGTGAAAG
7101 CAGGTGTGCT TTGTGACTTT TGGTGGTGGT TTGAAAAACT CAGAAAAGTT
7151 TAAAGAAGAA AGATAACTAG TATTCTCATT GTCCAGAATA TGATTTTTTA
7201 AATGTCTATA GAATATCACC ATCTGTAATT CTCCCGTAA TTTAAGTATT
7251 CAGTAGTTGT ATAAAACCTT TAAAATATAT ATATTGAGAA TTTTGTGTGA
7301 ATGAGATGAT GAGATAATCT TGTAGGATCA TTTAAAGATA AGAACTGAGG
7351 CCTGGCACAG TGGCTCATGC CTATAATCAC AGCACTTTGG GAGGCCCAGG
7401 CCGTAGATCA CCTGAGGTCA GGAGTTTGAG ACCAGCCTGG CCAACATGGC
7451 AAAACCTCTG CTCTACTAAG CATAGAAAAA TTAATTGGGT GTGGTCGTGC
7501 CTGCGTGTAG TCCCAGCTGC TTGGGAAGCT GAGGCGGGAG AATCTCTTGA
7551 ACCCTGGAGG TGGGCATTGC AGTGAGCTGA GATTGCGCCA CTGCCTCCA
7601 GCCTGGGCGA CAGAGCAAGA CTCTGTCTCA AAATAAAGTA AAATAAAGT
7651 AAGATAACAA CTGAAATTTT ACATTAATAA TTTTTTTGTA GCGACTGTGC
7701 CTCCTATGTT GTGCAGGCTG GTCTCAAACCT CCTGGCCTCA AGCGATCCTT
7751 CCAAAGCACT GGGTGGGCCA CCATGTCCAG CCTGAAATTT TGCATTAAAA
7801 AATTTCCCGC TTTTGGCTGG GCGAGGTGTC TCACGCCTGT AATAGCAGTT
7851 TGGGAGGCCG AGGCAGGCAG ATCACTTGAG GTCAAGTTCTA GACCGCCTG
7901 GCCAATGTGG TGAAACCTTG CCTCTACTAA AAACACCAAA TTAGCTAGGC
7951 GTGGTGGTGT GCGCTGTGAG TCCAAGCTA CTGAGGAGGC TGAGACAAGA
8001 GAATCGCTTG AATCTGGGAA AAAAGAGTTG CCGTGAGCCA AGATTGGCCA
8051 TGTCACTCCA GCCTGGGTGA CAGAGTGAGA TTCTGTCTCA AAAAAATAAA
8101 AAATAAAAAA TTCCCCTTTT AATCAAATTA AGTTAAATG AGGGATGTTA
8151 GACAGTTTTT AACCATCAAA TATTTTAGTT TAGTTTTTTT TTTTAAACGT
8201 TGTCTTAAAG ATGGAAGTGC TTCAAATCA AATCTTCTCT GCCAGTTCTC
8251 TACTTGGCTT CTTTTTTTTT CTTTTTGAGA TAGAGTCTCA CTTTGTCACT
8301 GGAGTGCCTT GCGGTGATCT CCGCTCACTG CAACCTCCGC CTTCCAGGTT
8351 TAAGTGATTG TTCCACCTCA GCCTCTCAAG TAGCTGGGAG TACAGGTGTG
8401 TGCCACCACA CCCGCTAAT TTTGTAGTT TTAGTAGAGA CAGGGTTTCA
8451 CTATGTTGGC CAGGCTGGCC TCAAACCTCT GACCTCGTGA TCCACCCACC
8501 TCAGCCAAAT TGCTGGGATT ACTTGTGTGA GCCACGCGCC TGGCTTCTAC
8551 TTGGCTTTTA AAGGGAATTT TGCTTTCTGA GTAATTTTAT TTCTCAGGTA
8601 TCTTGTGCTT TTTAATTCTG GAAGCAATCT TAATAATTTA TGTATGTGCC
8651 CTGTAATCCC AGCACTTTGG GAGGCCGAGG TGGGCGAATC ACGAGGTCTG
8701 GAGATCGAGA CCATCCTGGC TAACACGGTG AAACCCCATC TACTAAAAAT
8751 ACAAAAAAAT AGCTGGGCGT GGTGGCAGGC GCCTGTAGTC CCAGCTACTT
8801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN

FIGURE 3, page 3 of 21

1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379</
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

[REDACTED]

12601 GCTGGGATTA CAAGCACCCA CCACCATGCC CGGCTAATTT TGTGTATTTT
 12651 TAGTAGAGAC TGGGTTTCAC CATGTTGACC ACGCTGGTCT CGAACTCCTG
 12701 ACCTCAGGTG ATCTGCCTGC CTTGGCCTCC CAAAGTGCTG GGATTACAGG
 12751 TGTGAGCCAT CACACCAGGC CTCAAGAACT TTTTATTTTT GAGACAGGGT
 12801 CTCACCTCTG CACCCAGGCT GGAGTACAGT GGTGAGATCA TGGCTTACTG
 12851 CAGCCTGGAC TTCCCAGGCT CTGGTGATCC TCCCATCTCA GCCCCTGGAG
 12901 TAATTAGGAA TATAGACACA CACCCATGCC TGGCAGTTTT TGTATTTTTT
 12951 TTCTTTTTTC TCTTTTTTTG TAGAGACTGG GTTTCACATG TTGTATCAGG
 13001 CTGGTTTTGA ACTCCTGAGC TCAAGCAATC CTCACCTCTT GACCTCCCAA
 13051 CGTGCTGGGA TTACAGGCAT GAGCCACTGT ACCTGGCCTT TTCTACATTA
 13101 AAAACTTTTT ATTA AAAAAC CCAAATCTTC CTTGTGGTTG TATATACATA
 13151 TATACATAGG TACACACATG GAGAATTTTA CCTTGGAGGA AGGCTTGGTA
 13201 AAGAAAATAG CCCTTTGGGC CGGGTGCGGG GGCTGACGCC TGTAGTCCTA
 13251 GCACCTTTGG AGGCTGAGGT GGGCGGATTG CCTGAGCTCA GGAGTTCAAG
 13301 ACCAGCCTGG GCAACACAGT GAAACCTGT CTCTACTAAA ATACAAAAAA
 13351 TCAGCTGGGT GTGGCAGCAT GTGCCTGTAG TCCCAGCTAC TTGGGAGCCT
 13401 GAGGCAGGAG AACTGCTTGA ACCCGGGAGG CAGAGGTTGC AGTGAGCCGA
 13451 GATTGTGCTA CTGCACTTCA GCCTGCGCGA CAGAGCAAAA CTCTGTCTCA
 13501 AAAAAACAAA CAAACAAACA AAAAAGGAAA ATAGCCTTTC TCTATCATCA
 13551 GAGTATATTA AGAGTTGAGT TTTTTTTTCT GTTTTTTAAA ATTTTTGTTG
 13601 TTTATTTTAA ATTACAAAAC ATGGACTCTG CTTACAAATT AAGAAAATGA
 13651 CTCATGTTCA AACAAAGCATA ATCAATATAA CAGTTAATAC AAGTTAAATA
 13701 TTGTAATATG TTTACGGAAT AGCATGGCAA AATAGTGCAA AAGATTTGGG
 13751 GAAGGGGCTC ATAATTTCTG TTAACAGAAA GTTTTAGTTA TGTTGATTCA
 13801 ACTGGAGAGG AACAGAGCTC CCAGAAGGAC TCCAGAACAC TTGATGCTTG
 13851 TCTGAGTGGG GTCAGCAGCA CTGAGTTCCC ACCAGCCAGA AAGTTTGTGT
 13901 GTGTACATTA TTTCCCTTAA CTGCCACAAT AATCCCATGA AGAAAATGCC
 13951 CTAGTTTTTAC AAACAAGGAA ACAGAGGCAG AGAAGAGTTA AATGACTTGC
 14001 CCAAGGGCAT TCAAAGTAAG CAACTGAATT GGAATTTTAA CTCAAAGGCT
 14051 TGGATGTCCC ACTACAACAA ATAGGCTGTT TCTGCTTTAC TACATGTGCT
 14101 TACTTCTAAG AATTTAACAT TTTAGGCTGG TTGTGGTGGC TCACTCCTGT
 14151 AATCTCAGCA CTTTCGGAGG CTGAGGTGGG TAAATCACTT GAGCTCAGGA
 14201 GTTTGAGACC AACCTGGGCA ACATGGTAAA ACCTCATCTC TACCAAAAAA
 14251 AAAAAA AAAA CTAGCTGGAC GTGGTGGCAC GCGCCTGTGG TCCCAGCTAC
 14301 TCAGGAGGCT GAAGTAGGAG GATCGTTTGA GCCTGGGAGG TGGAGGTTGC
 14351 AGTGAGCCCA CATTGCATCA CTGCACTCTA GCCTAGGTGA CAGAGTGAGA
 14401 GCCTATCTCA CACACAAAAA AAAGAATTTA AAATTTTAGT CAAGTAATTA
 14451 GGCACATAA TTTTGTGGTC AGTTACTTTA CGAATTCATG GTTGGAGGCC
 14501 TGATGTGGTG GCTCATGCC TTAATCCCAG CACTTTGGGA GGCTGAGGCA
 14551 GGAGGATTGC TTAAGGCCAA GAGTTCAAAT CAGCCTGAGC AACCTAGTAA
 14601 GATCCCCCTT CTGCAAAAAA TTTAAAAATT AGCTGGGCAT GGTAGTGTGC
 14651 ACCTGTAGTC CCAACCACTT GGGAGGCTGA GGTGGGAGGA TTGCCTGAGG
 14701 CCAGGAGTTT GAGACCTGGG CAGCATATGA AGACCCTGTC TCTAAAAAAC
 14751 TAAAAATAAA AAATAGCCAG GTGTGGTTGG TGTGCTTG TGTCAGCTA
 14801 CTCAAGAGGC TGAGGCAAGA GGGTTGCTTG AGCCCAGAAG TTGGAGGCTG
 14851 CCGTGAAC TGATTGCACC ACTGCACTTC AGCCTGGGTG ACATAGCAAG
 14901 ACCCTGTCTC TGTGGTGGTG GTGGGTGGGG GTGGGGGAAG GGATTTAAGA
 14951 AGGGTTTGTG AGGTATGTAT TATTTATAAA TGGGCTTTTA ACTTTACCTT
 15001 TCACATCTTG GGTGAAATT AATTGTATCC ATTCTCAGTT TTTCTGTCTT
 15051 GCTATATATT TAACTTGGA GACTTAGAGG TCATGGATGT CTTTCTATGA
 15101 AAAGCAAATG AAGCAGAGGG CTGCCCTTCTC TTGCTGTAGA GGGCACACTT
 15151 GCTGCAGAGC ATGTTACTGT TTTATGCATT GCTAGGCTTT GGGAGTTGTG
 15201 ACTTGTATGA TCATAGTACT TACAATATT AGTTGGCAAT TTTTAACTT
 15251 TAACTTTAGA TTATATATGT AAACCTCTGT GTTCCTTTGT CACTGATAAT
 15301 CTGAACAGAA GCCTTGGATA AATAATTTTG AAGTTTTTGT CTGAACCTCT
 15351 GAAATTTGTA TTGTTATCTC ATGGTTTTGC TGGGAGGAAG GAGAAATAAC
 15401 AATGGCCACT TACTGTGCTT CTGTATGTGC CAGACAGTAT GTGCTAGATG
 15451 TTTCAGAAAC GTGATTTGTA ATCCTGACAA GAAGCCTAAT TGGGTGGTAG
 15501 TGGGTGCTAA TTGAACCTTA TAGATGAGGA AATTGAGGCT CATGGTGGTA
 15551 AGTGAATAAC TTGCACCAAG ATCCTATGGC TGGTATGCAG TAGAGCCTCA
 15601 ATTCAAGTAC GGGTCTTCCA GGTCCAAACC CATGCAGGCT TTGAGAGGTA
 15651 AGGAGGTAGA GAACGTTGAC ACCCCTTCT TGGTGTGTTT TTCAGCAAAT
 15701 ACTTGTATGC ATATTAAAGA CTGTCTACCC TTTTGTCTATC TTGTGTCACT

FIGURE 3, page 5 of 21

15751
15801
15851
15901
15951
16001
16051
16101
16151
16201
16251
16301
16351
16401
16451
16501
16551
16601
16651
16701
16751
16801
16851
16901
16951
17001
17051
17101
17151
17201
17251
17301
17351
17401
17451
17501
17551
17601
17651
17701
17751
17801
17851
17901
17951
18001
18051
18101
18151
18201
18251
18301
18351
18401
18451
18501
18551
18601
18651
18701
18751
18801
18851

TGCTGCTTCC TTTGGTACTA CCCAAATTTTCTTCAGCATT TCAGCTTTGA
ATTTTTATTT TTATTTTATT TAATTTATTT ATTTTTTTGA GATGGAGTCT
CACTCTGTTG TCCAGGCTGG AGTGCAGTGG CGTGATATCA GCTCACTGCA
ACCTCTGCCT CACAGGTTCA AGCAATTCTT CCTGCCTCAG CCTCCTTAGT
AGCTGGGACT GGAGGTGCCC ACCACCACGC CCAACTAATT TTTGTATTTT
TAGTAGAGAT AGGGTTTTAC CTTGTTGGCC AGGCTGGTTT TGAACCTTG
GCCTCAAGTG ATCCACCCAC CTCGGCCTCC CAAATGCTG GGATTACAGG
CATGAGCCAC TGCACCTGGC CAGCTTTGAA TTTTATAGAA ACTGTTCTAA
ACAGAACTAT ATTGGAACCT GGAAAATTAA TCTATTGTCT CTAATACCA
AAGAAAAACA TGTAATTTTA GTGGTTGATT ATGGGAACAA TTTTTTTTAA
GATGGTTCAT CTGAATGGGA AGCATTTTTTT TTTTAATTGC TTGACTATTT
CTTTAAATTT GGAGAAAAGA CCATTGCCCT CTCAGATTTC TGGTAATTGG
TCACATTGAT CATTATATTT GACTGACAGG CTGCTTTGTC CACAGCTGAA
GGATTGTTTA ATTTTTTTTA AATTATAAGA GTAATATGTG CTCACTGTAA
AATTCACAGT ACAGAAGCAT ATGAACTAAC TAAAAGTTCT TACCTCTTGT
CTCCAGCAAG GAGTAAGTGT TTCAACCTGA AGGTTGGTTT TGAATTGTGT
TCTGTGGAGC GTACTTAAAG TGAGTGAAGA AGAAAAATTT ATGTCAATCA
TATCATTGCT AGCTGAAGTT TTTATTGTTT CACCCCCTAA AGGTTATTAA
AATAGTATGT AGTTTAGTAG TCTTGATAAT TTTCCCTTAA GATTTATTGG
CCAGTATATC AGGATTTTGT TTTAAATTTG ATATGTGAGC TTAGTTTAT
GCTATTTTCA AATAAGACAT TTAGAAGAAG ATAAAAATAC ATTCCTGTCT
TAGTCTGTTT TCTGCTGCTA TAACAGAATA GCACAGACTG GGTAAATTTAT
AAACAGTAGA AGTTTATTTG GCCTGTGGTT CTGGAGGCTG GGAACCTCAA
GAGCATGGTT CTGCCCTTTG TGCTGTGTTA TCATATGGTG GAAGGTGGAA
AGGCAAGTGG GTATGTCAAG ACAGAGAGCA AGAAGGGGCT TGAACCTCACT
TTTATAACAG AGTGACTCCA GAGATAGCTA ACCCACTTTT GAGAGAATGC
ATTAATCCAT TCATGAGGGC AGAGCCCTTG TGACCTAATC ACCTCTCATT
AGGCTCTGCA TCCTTAAACT GGTTTTTTTTT TGTTTTTTTTT TTTTGAGACG
GAGTCTGCT CTGTTGCCCA GGCCGACTG CGGACTGCAG TGGCGCAATC
TCGGCTCACT GCAAGCTCCG CCTCCCGGGT TCACGCCATT CTCCTGCCTC
AGCCTCCCGA GTAGCTGGGA CTACAGGCGC CCGCCACCGT GCCCGCTAA
TTTTTTGTAT TTTTTTAGTA GAGACGGGGT TTCACCTTGT TAGCCAGGAT
GGTCTCGATC TCCTGACCTC ATGATCCACC CGCCTCGGCC TCCCAAAGTG
CTGGGATTAC AGGCGTGAGC CACCGCGCCC GGCCCCCTT AAAGTGTGT
ATTGGGGATT AAGTATCTAA CACAGGAAC TTTGAGGATA CATTTAAACC
ATAAGAAATC CTGTCATGCA AATGAATCCA TTCTAGATGA AAGAGAATGA
ATTTAGTTTC CATTGAACCT TATAAATAGG CCTTTTCTAA GGTACTTACA
GCTGATATTA TAAAATTTAT ATTTGTTTTT ATAAATTTGT ATTTGTATTT
CTGTTTGATC AAATACAATT ATACACTATA GTTCTCTGCT GTTAGATTTT
TTTTCTTCCT TAGCATGTTT CCAAAGGGTG GAATGTTGAA AGTTGGGTTA
ATGTCAATCA GCTTTCCTTT GTAAAGTGT CATTGACATG TGAACCTTGT
CTGAGAATCT AAATTTTATT TCATGAAAGA AGAAAACAGT ATATTCTCAT
TTAACCAGAA ATTTAACTTC ATATACTTGT GGCTGTATTG GGAGTATGCC
ATTGCTGTCT GTTTACAACC TGACCTACTC TACCTACTTA GAAGTAATTT
GTGTTATGAT AGGTGTGCTG TGCTGACATA TGCTGAACAT ATTTGTAAGG
GTGTTAAGTC ATTGAATAAA ACGCTTTTCT CCTCCTTTCA AATAACATTT
TTTATTCTCG GTTATAAAAG TCATACAAGC TTAGTGCAGG TTGTTAAAAA
GGTATAAAGA AGAAACCGTC AATCCATTAT AATCCTACAG TTTAGACTTC
CTGCTCCAGC CTCTCAGAGT GCTGAGATGA GCTAGCCATG CCCAGCCCCT
CAAAAGATTT TTTAAAAAAC AAAAATGAGG TTATACTTTA AAAAATCTA
TATTCCTTTC ACATAACAGT GTTATTTTGG AGGTTTTAGA ATTTCCAGTA
GCATTTTAGA TTCAGAAACA AGCTGATTCA TCCTCTACTT TGTACTTTAG
GCAAGAAAAG AATTTTACCT AAATAGAATT TTGAACTGAA AATCTGTTTT
TCTAACTTTT TATTTAAAGA ATATTGTTCC ATGCTTTCAC AGTAGTGACT
TTTAATTTTT ATATTTTTTA TTTTATTTAT TTAGAGATGG GGGTCTCACT
CTTGTTGCCT AGGCTAGAGT GAGTGCAATG GTTCTATTCC TAGCTCACTG
CAACCTTGAA CTCCTGGGCT CAAGTTACCC TCCTGCCTCA GCCTTCTAAG
TAGCTGGGAC TACAGGTGTG CACCACTGCA CCAGGCTTTT TTTAAAGGCA
TAGAAAATGG TAGTGCTTGC ATACAAAAAT GGCGTAGGTA CATACATCAG
CGGACATCAA GACTATGTTT AGATCATAAA TGTACATATA TGTACCGATG
CCATTTTTCG ACGCAAACAA ATAATGGAAA TTGAACTCTA AACTGAAATT
TGAAACAAGG GTTCTGGGGT GGGCCCTCTT GCTGATTTGT AATTGAATGT
ATAGTTCAAT TTTTCCCAT CTGTTAAGCA AAAGACAATT CTAATGTTAG

FIGURE 3, page 6 of 21

22051 GAACTACAAA AAGCTAGAGG AAGTTCTGAA CTGGAAACAG TGGATAGGAT
22101 TTACTAGAAT AATTTACGAG GGTGACAATT GTAAATCTTC ATAGGTTTCT
22151 TTTTCTTCTT TCTCTTTTTT TTTTCTTTGA GATGGAGTCT CGCTCTGTTG
22201 CCCAGGCTGG AGTGCAATGG CGCAGTCTCT CCTCACTGCA ACCTCCGCCT
22251 CCTGGGTCCA GGTGATTCTC CTGCCTTAGC CACCCAAGTA GCTGGGATTA
22301 CAGGCATCTG CCACCATGCT GAGCTAATTT TTGTATTTTT TTTTCTAGTA
22351 GAGACGGGGT TTCACCATGT TGGTCAGGCT GGTCTTGAAC TCCTGACCTC
22401 AGGTAATCCA CCCACCTTGG CCTCCCAAAG TGCTGGGATT ACAGGTGTGA
22451 GCCACCGGCG CCAGCCAAAT TTTTATTGGT TTCTAAACTA CCGTAATTTA
22501 GTTTTTTCCA CTTAAGTCAA AATTATATTA TTGTAGGATA AAAACTTAGT
22551 GATCCAAATT CATGAGGAAT GAAGAATAAA TACATTTAAA GTCTTACCAT
22601 TTGCTAAATT AGTCTTGGCT CTTTGTACCA AAATTTCTGT CTTGTGCTCT
22651 GTAATTTTAT ATTTGTATAT TTTCTATCAA CATTTTTACT GTGTGGTGT
22701 TATGTAAATTA TAAAAACGTT TTAAGCAAAA CTCAGAACAA TGAATTCTCA
22751 CGAATATTCA GTATATTTAC AGTTGAGAAA TAACTACTT CTGTAGTAGG
22801 TAATTTAAAA TGTCCTCAATG CAAGTTAACG TGCTACTGAT CACGCTATTC
22851 AGGTGTGTGT CTTTGATAAG GGGAGGTGGG GAAGTTTGTG GGTTTGATTT
22901 TATTTGCCCT TCTCATGTGA CTGTTGTCAT GTTAGTAAAC AAATGGTTTG
22951 CGAGAGAACC AGTAGTCTTT TGCAAAGATT GTCTTATACA GAGCACTCAA
23001 TTCTTCATAT TATTTATAAT GGCTTTAATT TAAGCCTTAA ATTATTAGAA
23051 ACTCATAAAT AATTTTTTTA TTTGTTTTTT TGAGATGGAG TTTCGCCCTT
23101 ATGTGCCAGG CTGAAGTACA ATGATGTGAT CTTGACTCAC TGCAACCTCC
23151 GCCTCTCGGG TTCAAGTGAT TCTCCTGCCT TTGCCTCCCA AGTAGCTGGG
23201 ATTACAGGCA TGCCTACCA TGCTGGCTA ATTTTGTATT TTTAGTAAAG
23251 ACAGGATTGC ACCATGTTGG CCAGGCTGGT CTCGAACCTC CAACCTCAGG
23301 TGATCCACCT GCTTCGGCCT CCCAGAGTGC TGGGATTACA GGCTCACTGA
23351 GCCACTGTGC CCAGCCATAA TGCGTTAAAA TAAGAGTGT ATATTTGTAA
23401 AACTTAAAAA AATGTAGTGG TTGAAAAAGG TAATTTAAAA AGAATTGACT
23451 ATTAATTTCT TGAAACCATA ATGTAACCTG TAGTGCAATT AGGAAACCTT
23501 CATGTTTCTT TCTTTCTTTC TTTTCTTTTT TTTTGTAGAT GGAGTTTGTG
23551 CTTTGTGCTC TAGGCTGGAG TGTGTGATGT CAGCGCACTG CAACCTCTGC
23601 CTCTGGGTT CAAGCAATTC TCCTGCCTCA GCCTCCCGAG TAGCTGGGAT
23651 TACAGGCGCC TGCCACCACA CCCAGCTAAT TTTTGTATTT TTAGTAGAGG
23701 CGGGGTTTCA TCGTGTGGC CTGGCTGGTC TCGAACTCCT GACCTCAGGT
23751 GATCCACTGC ACCTGGCCCC CGTTCATGTC TTTTAAAGCT TTATGGTTGC
23801 TCTGAAATAG AGTTGTGAT TTTTCTTTTT TTTTGTAGAC TCCTCTTTTG
23851 CCCGTGCTGG AGTGCAGTGG TGTGATCTGA GCTCACTGCA ACCTCCACCT
23901 CCTGAGTTCA AGCAATTCTC ATGGGTCAGC CTCTCAAGTA GCTGAGATTA
23951 AAGCTGCCCA CCACCATGCC TAGCTAATTT TAGTATTTTT AGTAGAGATG
24001 GGGTTTCAAC GTATTGGCCA GGGTGGTCTG GAACCTCTGA CCTCAGCAT
24051 GAGCCACTAC GCCTAGCCTG GGTGTTGAT CTTTAAGGTG ATACTTCAGG
24101 CAACATCTGA GGCCAGTAC AGTCCTTTAC TTCAACTGGC TCCAGTACAG
24151 CAAATTCAGG GAATGTTTTT GAGTGTTTAC TGATGCTG CCGTGGAGTT
24201 CAGGAGATT GGTACATTGA GTCCAGTTGT TGTGTTGAAA CTCTGTGTTA
24251 AAAACCTCCC TACTAAGTCC CAGCTACTCA GGAGGCTGAG GCCTGAGAAT
24301 CACTTGAACA CCTGGAGGCA GAGGTGTCAG TGAATCGAGA TCGAGCCACT
24351 GCACTCCAGC CTGGGCGACA GAGTGAGACT GTCTAACAA AAAACAACA
24401 CCCCCAAAA AACCAACCTA CTATGGTAGT ATCAATGCTG TGATAGTCTT
24451 CCTTCTTCA TACAGGTAAA TTCTTAACAT ATACTCATG TTAATGTTC
24501 GTGTTCACTA TTCTTAAGAG TATTTGGGGC CAGGCACGGT GGCTCATGCC
24551 TGTACTCCA GCACTTTGGG AGGCTGAGGT GAGCAGATTA CCTGAGTTA
24601 GGAGCTTGAG AACAGCCTCC AACATGATGA AACTCCCGTC TTTACTAGAA
24651 ATACAAAAAT TAGCTGGGTG TGTTAGCACA TGCTGTGTA CCCAGCTACT
24701 TCAGAGGCTG AGGCAGGAGA ATTGCTTGAA CCTGGGAGGT GGAGGCTGCA
24751 GTGACCTGAG ATTGCTTCAC TGCACTCCAG CCTGGGCAAC AGAGCGAGAC
24801 TCTTGTCTCA AAACAAACAA ACAAAAAAAG AATATTTGGG GCCAGGCATG
24851 GTGGCTCACA CCTGTAGTCC CAGCACTTTG GGAGGCCAAG GTGGGTGGAT
24901 CACTTGAGAT CAGGAGTTGG AGACCAGCCC GACCAACATG GCTAAATCCC
24951 GTCTCTACTA AAAGTACAAA AATTAGCTTG AGCAACAGAG CAAGACTCTG
25001 TCTCAAAAAA AGAAAGAAGA ATATTTGGTT TAATTAAGAA GGAACCTTAT
25051 CAATAGTAGT AAAGTCAGCC AGCTGAACCT CCAAGTACAA ATTGTTGGTA
25101 TTAGGTATCA ATCATTTATT AAGGATAATA TTCTACAATA GCGATCTTTT
25151 TAAAAATTTT AAAATCTCAA ACTGGAAAGG ATGTCTAGTT CATTCTATGC

FIGURE 3, page 8 of 21

25201 TTCAGTCCCC TCTTCTGATT TACTTGTTTA GAAGATTTTT GTTTCCTTCT
25251 CTGACTTCTA TTTTGTGCT GACTGGCACT TGGGATTTTT AAAAAATTAT
25301 TTTCTCTATA TATAATTAAA GACAATAAGT ATAACAATAA GTATAATATG
25351 GTAATTTGCT AAAACCCAAA CAATGTTTTA AGTAATGCAT ATCATTATGT
25401 AAACCTACGT AATAGTTGAA TATTCACAAA GATAATCGCT TATAGAAGTT
25451 TTATATCCTC TCTTCTTTGG CAGTGCAATT AAAACAAAAA AAATAAGTTT
25501 TATGTCTTGT TTACATGTAA ATAATTTTAA TCTAAATTGT GACGTGGTTT
25551 TCACTTTAGC ATATTTTGA AAGTAAATCA AAAAGGACAA AATACAAAAAT
25601 CATGTATATC TTCTACAAAA ACGATATATA AATTCTAAGG TTTTGTCTCT
25651 TTTGAAATTG CTTAAAAAGAA TGCATAGAAC TGGTGTCTGA GTTGGGAAGG
25701 ATCTATGAGG GATTTCCTTG GAGACCGTGG GTGAATAATA ATGTTGTCTT
25751 AGTTCCATGA AGGAATCTCT GGGGATAGTT TTTGAGTTAG GCCTGGCAAT
25801 GTTAGAGATA CATAAAGAGA GCCTTGTTTT ATCACTGGGT GCGGTGGCTC
25851 ACACCTGTAA TTCCAGCACT TTGGGAGGCT GAGGCGGGCA GATCATGAGG
25901 TCAGGAGATC GAGACCATCC TGGCCAAACAC GGTGAAACCC GTGTCTACTA
25951 AAAATACAAA AATTAGCTGG GCGTGGTGGC GCATGCCAT AATCCAGCT
26001 ACTCGGGAGG CTGAGGCAGG AGAATCACTT GAACCAAGGA GTTGGAGGTT
26051 GCAGTGAGCC GAGATCGCGC CACTGCACTC CAGCCTGGGT GACAGAGCAA
26101 GACTCCGTCT CAAAAAATAA AAGCTTGGTT TTCAATGGTT CTGAAAAATG
26151 CTTTAATACA AGTGTAGAGT GTTAGTCAAG TTTTGCACCT GGATAAACAG
26201 CCTGTGAATT TATCACATTT CTAGTTTATA ATATGGGCTT TCAGAAGTTA
26251 TATGAACATT GTTTTGACGG GAGAATTCAA GCTGGATGCT AGAGAAGGAT
26301 CGTGAGAACC CCTTCATTGG AGGAGTGCTA TGAAATTATT TGATCTTGGA
26351 ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTGTAGAC AGAGTTTCGT
26401 TCTTATTGCC CAGGCTGGAG CTGGAATGCA GTGGCACGAT CTCGGCTCAC
26451 TGCAACCTCT GCCTCCTGGG TTCAAGCAAT TCTTCTGCCT CAGCCTACCA
26501 GGTAGCTGGG ATTACAGGCA TGCGCAACCA TGCCAGCTA ATTTTTGTAT
26551 TTTTAATGGA GACGGGGTTT CACCATGTTG GTCAGGCTGG TCTTGAACCTC
26601 CTGACCTCAA GTGAACCTGCC TGCCTCAGCC TCCCAAAGTG TTGGGATTAC
26651 AGGTGTGAGC CACTGCGCCT GGCCTGATCT TAGAATTGGA AGGAGAGACT
26701 AATATTTTCA TGGCAAAAAC AATGAAAAGT TACCTTTCTG TATCTAATA
26751 CTATAGAGGA GTGGGATTTA TTTAGAATGT TTTAAGTATC TTGGGCAGTC
26801 CAAGAGTGCG TATCACTTAT TTTTCTTTTC CTCTTTCTT TTTAAGTGGA
26851 AGTTCACTGA TGTAGAGAT CATAGGTGGC ATTGCCCTACT TTTTACATAA
26901 TTTTATCATG TTTAGTGATC TGTGAGAGG GCTGTGGCTG TTTGCAGTTT
26951 TGGCTTAAGC CATGCATGGG CTTTATAGGA GATGTAGTCT TCACAGTGAG
27001 TTGTTATTTG TAGCTGTGTT TTTGTTTTTG TATAGCTTAT AGCAATGCAG
27051 TGTGCTTTTT ATTAACATCA TTTTCTTTTT CTTTTTGCAG TGATTATTTA
27101 TTCAAGTTAC TTCTGATTGG CGACTCAGGG GTTGGAAAGT CTGCTCTCT
27151 TCTTAGGTTT GCAGTAAGTT GAAATTGAAA TGTCTTTACA ATTAATGGTA
27201 CAATTAATGC TATGTATGTT TTCTAGGTAG ATAAATTAAG ACAGTTTTAT
27251 TCAGAATAAG TTAATCTTTC CAGAATTTAT ATATTTAAAG ACTCCAAATA
27301 TACATCCCCA GTGGTATCTT GGACTGTAA ATAGAAAAAT ATTGTTGCTC
27351 TTAAGAGAAA TTCAGTGAAG TCTGGTTATA AAGTCAGAA GTCTAATACT
27401 TTTGGTCAGA GTCAAAACAGC AGTTCCAATA TAGGCAGCAA GTTAAAGGGG
27451 TAGTTGGTGG CCTGTGTTGA AAGCGACTTG ATGAAAATAA ATCTTTAAAT
27501 TAAACTTTAG TAGAATAAAA AGAAAAAGCA GAGCCAGGTG ACGCAGTGGA
27551 TCATGCCCTGC AGTCTCAGCT ACTCAGGGTG CTGAGGGTGG AAGGATCACT
27601 TGAGTCTAGG AGTTTTGAGA CCAACCTGGA CAACATAGCA TGACTCTGTC
27651 TCTGAAAAAA AAAGTTAATA AAAGAAAAAG TAGGGTCTTG GACAACTTC
27701 GTTGGCCAAT GGCATAGTTC TAAATGCTGA AGCTGACAGA TAAAGGACTT
27751 TTGACTTAAC AGAATCCACA GTGTCCTTCA TAGTCTTTAT CAACTACCTT
27801 TAAATTTAGC ATGTTTCCTG GCCAGGTGCG GTGGCTCACG CCTGTAATCC
27851 CAGCACTTTG GGAGGCCGAG ACGGCGGGAT CACAAGGTCA AGAGATTGAG
27901 ACCATCCTGG CTAACACGGT GAAACCCCGT CTCTACTAAA AATACAAAAA
27951 ATCAGCTGGG TGTGGTGCCA CACGCCTGTA GTCCCAGCTA CTCGGGAGGC
28001 TGAGGCAGGA GAATCGCTTG AACCCAGGAG GCGGAGGTTG CAGTGAGCTG
28051 AGATGGTGCC ACTGCACTCC AGCCTGGCAA CAGAGCAAGA CTGTCTCAAA
28101 AAAAAAGAA AAAAAATAA AAAACAAAT AGCATGTTTC CCTTCTAGAG
28151 ATCATTTGTT CTGAGAGCAT GGACCAAGA CTCCTGGGGG TTACCAAGAC
28201 CCTCTCAGGT AGCCCATGAG GTCAAAATAT CCTAATAATA CTAAGATGTT
28251 AGTATTTGTA AGGAAATATT TACTTGGTAA TAATACTAAT ATAAAGATG
28301 TTTGCGTTTT TCAGTGATGA CATTGGCTCT GGTACAAAAG CATGTGGGTA

FIGURE 3, page 9 of 21

Figure 3, page 10 of 21

28351 AAATTGCTGC TGGCTTGGTA CACATCAAGG CAGCGCTAAG CTCCAAATTG
28401 TACTCATGGT GATGGCATT TTTACCTCTG TGCCCTCACA GGAACAAAAA
28451 CAAGCCGTGC CATTTTATT GAAGATTGTC CTTGACAAAA CAGTTAAAAA
28501 GATTAATTTT TGAAAAATGT TGATCCATGA GTATTCCTTT AAAAAATTTT
28551 GTGAAGAAAT GGGAAAGTCA CATAAAACAA TGTTTTTTTT TTGTTTTTTT
28601 TTTTTTTTTT TTTTGAGACA GATTCTGGCT GTGTTGCCAA GGCTAGAGTG
28651 CAGTGGCGTC TGGCTCCCAG GCTCAAGCTG TTCTCCCACT TCAGCCTCCC
28701 AAGTGGCTGG GACCTCCCAA GTGGATGCGC CATCATGCCT GGCTGATTTT
28751 TGTATTTTTT TGAGTGACA AGGTCTCACT GTGTTGCACA GGCTGGTCTC
28801 AAACCTCTGA GCTCAAGCGA TGCATGTGCC TCAGCCTCCC AAAGTGCTGG
28851 AGAAAGCACT TTTTACTGCA TACTGGCTAG TGTGTTGGTT ATTTTGAGAG
28901 AAAGAAAAGC ATTTGTAGTT TTTTGAGTTG TAAGCTGAGC TAACTGCTTT
28951 ATTTTCTTCT GTGGAACACC ATTTCTTTTT TTTTTTTTGA GATGGAATAT
29001 TGCTTTGTTG CCCAGGCTGG AGTGCACTGG CACAATCTCG GCTCACTGCA
29051 ACCTCCGCTT CTCGGGTTC ACGAATTCTT CTGCCGTAGC CTCCCAAGTA
29101 GCTGGGATTA TAGGCACCTG CCACCAAGCC CAGCTAGTTT TTGTATTTTT
29151 AGTAGAGATG GGGTTTCACC ATGTTGGCCA GGCTGGTCTC GAACTCCTGA
29201 CTTCTGATC CGCTTGTCTC AGCCTCCCAA AGTGCTGGGA TTACAGGCGT
29251 GAACTACTGC ACCTGGACAT TTTTTTTTTT TTTTAACTT GAAAGAACAG
29301 CTAACAGACA GATTAGAACA GAATTGGCTA TTTGACAGAT TTTCTCAGAT
29351 GAACTGTGAT AGTCATTTC AAGGAAGTAG CTGCAAGCAT TTGTTGGCTG
29401 AAATAAAATT TAAGTTTATC ATGGAAAAAT AGAATTTGAA AAAACTTAGA
29451 GTTTACCACT TGACAGTATC CTAATATCAT ATGACTTTTC TGATGAGTGC
29501 CGATATTAAT GAAGGTTATT TAAAAAATAT TAAATAATGT ATAATCTTTT
29551 TTATATAACA GTTAAAAATA AAACCATGAG TACTAGAATA AAACATAGGT
29601 GGCTCTTAA TCTTGGTTTG TGAAGGTATT TTTTAAAAATA AGAAAAAAGC
29651 AAGAAATCAC TGCTAAATTT GACTATTAAA ATTAATTTAT CACAGGCACA
29701 AAAATGTTAG AAAACTAATG GCAATAGCAA ATATATATAT ATGAGGATTG
29751 GTATTCTCAA CATATAAAGC ACATTTGCAC ATCAACAAGA AAAGAATATT
29801 TCTCCTAATG GAAATAGTGG CAAATACATG AGCAGTCAGT TGAAAAAAGA
29851 AGTAATACAA ATTGCTGGCT GGGTGTGGGT GGGGTACAGC CTGTAATCCC
29901 AGCATTTAGA GGCTGAGGCT GCGGATCAT CTGAGGTCAG GAGTTCGAGA
29951 CCAGCCTGAC CAACATGGAG AAACCTGTG TCTACTAAAA ATACAAAATT
30001 AGCCGGATGT GGTGGCGCAT GCCTGTAATC CCAGCTACTT GGGAGGCTGA
30051 GGCAGGAGAA TTGCTTGAAC CCAGGAGGCG GAGGTGTGG TGAGTCGAGA
30101 TCGCACCATT GCACTCCAGC CTGGGCAACA AGAGCGAAAC TCCATCTCAA
30151 AAAAAAAAAA AAAAAAAAAA AAAAGGAAGT AATACAAAT GCCAATAAAT
30201 ATGGAAAAAA AAAAAAGGCTC AACTTTATTT GTAATTAAAG GCCTTTAAGT
30251 TAAACTTAGG TGTCATTAA TTTTATTAA ATTGGCAAT ATTAAATTA
30301 AGCATAATTC TTAAGCAACT CTCGGTAGGT GGAAGAATC TAGCTGTAGC
30351 CTCAGGTGTT TGTGCCTCAA GGAACCCCT CTCTGGGATG TCCATTGCTT
30401 GAAGTCAAAG GTTTTCCAAT AATACCTGGA AACTATTTT AAAATGCTGA
30451 TCCCATATAC CTCAAAATAT TAATAGAGAC AATCGTGAG ACTATAATAA
30501 AGAAATGTGC AATAAGCTCT GGGGCGACAG AGGGAAGAAT CTATTGGCTG
30551 AGGAGTTGAA GAAATTGTTT GGACACTCAG TATTGCCTGA GCTCAAAACT
30601 GAAGGATGAA TAAATGCCAC ATGACCTTGG GGCTGGGGAG TAAAGTAGGT
30651 TATGCAGAGA GAGATAACTG AGGCTTTTGG GCAGACGAAT AGTAACGGCT
30701 CAGGCATGGG AGTAAAGGTC ATTTAGAGAT TTACAAGAAT TCAGCATTTT
30751 TTTCTTTTTC TTTTTTTTTT TTGAGATGGA GTCTAGCTCT GTCATCCAGG
30801 CTGGAGTACA GTGGCATGAT CTCAGCTCAC TATAACTCCC ACCTCCCGGG
30851 TTCAAGTGAT TCTCATGCCT CAGCCTCCCG AGTAGCTGGT ATTACAGGCG
30901 TGTACTACTG TGCCCTGGCTA ATTTTGTAT TTTTAGTAGA GATGGGGTTT
30951 CACCATGTTG GTCAGGCTGG TCTCCAACCT CTGAGCTCAA GTGATATGTG
31001 CACCTCTGCT CCCCAAAGTG CTGGGATTAC AGGCGTGAGC CACTGTACCC
31051 GGCCAAGAAT TCAGTATTTT TATCCAAGTA CCTGGGGGAT AGATGTGCTA
31101 CATGAATATT TATTGCATTC ATTTTGTCT CTGCATTTT TTTTTTTTTT
31151 TTGGTTTGAG ATGGAGTCTC GCTCTGTGCG CCAGGCTGGA GTGCAGTCGT
31201 GCAATCTCGG CTCAGTGCAG CCTCCACCTC ATGGGTTCAA GCGATTCTCC
31251 ATCTTGCTCT CCTGACTAGC TAGGTTTACA GGCGTGTGCC ATCACACCCA
31301 CTAATTTTTT GTATTTTATG TAGAGACAGG GTTTCACCAT GTTGGCCAGG
31351 CTGGTCTTGA ACTCCTGATC TAAAGTGAGC CTCCCACCTT GGCCTCCCAA
31401 AGTGCTGGGA TTACATATGT GAGCCACTGC GCCTGGCCTC TATATACTTC
31451 TATAGTACCT GATACTTATT AGGCACTCAA TTACAACATA ACTTTTTTTT

FIGURE 3, page 10 of 21

31501 TTTTTTTTTT TTTTGAGACA GAGACATGCC TTGTCGCCTG GGCTGGAGTG
31551 CAGTGGCACA GTCTCGGCTC ACTGCAACCT TCACCTCCCG GGTTC AAGTG
31601 ATTCTCCTTC CTCAGCCTCC CGGGTAGCTG GGATTACAGG CGCCCCCAC
31651 CACGTCAGC TAATTTTTTG TATTTTAAAT AGAGATGAGG TTTCAACATC
31701 TTGGCCAGGC TGATCTCAA CTCCTGACCT TGTGATCCAC TCACCTTGGC
31751 CTCCCAAAGT GCTGGTATTA CAGGTGTGAG CCATCATGCC CGGCCCATAT
31801 TTCTAAAAAC ATTTTCTTAT AAAATGACAT TGCCATTATC AACCTGCAAA
31851 ATACATTTC ATTTGGTTGT TTTCTTGCTT AGTCTTTTAA TCTAGAGTTT
31901 TATACCTTAT CTTTTTTAT TATATATTTT TTATGTCATT GACTTTTTGTC
31951 AGAAACTGAA GCACTTGTC TGTAGATTGT CCAATATTCT AGATTTGTCA
32001 TTTTGTTC TTTGATGTC CTTATGCTTA TTTGTTTGT CCTCTTCTG
32051 TAATTAGAAG ACCTAGAACT GCACTATCCT TAGAGTAGCT ACTAGCTCTA
32101 TGTAGCTATT TAAATTTAAA TTAATTAATA TTGAAAAAGT TTGGTGGCTC
32151 ACACCTGTAA TCCCAGCACT TTGGGAGGCC AAGGTGGGAG GATTGCTTGA
32201 GTGCAGGAGT TCAAGGCTTC AGTAAGCTAC GATTGTACTC TAGCCTGGGA
32251 GACATCAAGA CCCTGTCCCT TTAAGGGGGA AAAATAATTG AAAAAATCAA
32301 AAACCTAGTT TCCTTGTTTC ACAAGCTGCA TAGGGCTAAT GGCTACCATA
32351 TTGGCTAGCA CAGCTTATAG AACCTTTCCA TTGTCACAGA AAGTTCGTGT
32401 TGGCAGTGCC GTTCTCATTA GACCTGATTC GATTAAGGTC CATCTTTGTT
32451 GACAGAGTAC TTCTTAGGTG GTGCTTTGTG GTTCATATGA TGATAGCCTG
32501 GTCTGTTTAT TCATATATCT TTTACGAGA AATATTTTTT TTCCATTCTG
32551 AATAAAATTT CATGGCAGGT ACTTGCAAGA AGCAGTTATA ATTTTAAAGT
32601 TTAACATTAG GTTAAAAAAT TGACAGGAAA CATATATTCA CAGGTAACAA
32651 TTGTACACAA ATGTTTCATGG CAGCATTATT CATAATAGCC AAGAAGTGGA
32701 AACACCCCAA ATCAATTTAT GAATGGATAA AATGTTGTAT ATTTGTAGTA
32751 CATGTAATAT TATTAGCCA ATAAATGGG CCAGGCATGG TGGCTCACAC
32801 CTGTAATCCC AGCACTTTGA GAGGCTCAGG CAGGGGGATC ACTAGAGGTC
32851 AGGAGTTTGA GACCAGCCTG ACCATCATCA CGAAACCCTG TCTCTACTAA
32901 ACGTACAAAA ATTAGGCAGG CGTGGTGATG CACGCCTGTA GTCCTACTA
32951 CTCAGGTGGC TGAGTCATGA GGATTGCTTG GACCCCGGGA GACAGAGGTT
33001 GCAGTGAGCT GAGATCATGA CACTGCACCT CAGCATGGGC AACAGAGCAA
33051 CATCTGCCT CAAAAAATAA AAAAAAATAA AAAAGAAGTA CTGTTACATG
33101 GTACAACATG GATGAACCTT GAAACATTTC TGCTAAATGA AGGAAGACAG
33151 ACACAGAGGG CCACATATTT TATGATTCCA TTTATACGAA ATGTCCAAAA
33201 TTGGCAAATC TAAAGAGAAA GTAGATTAGT GGTGGCCAGG GAGTGAAGAC
33251 GGGTTCTTTC TGGAGTGAAG AAAATGTCCT GGAATTCGTG GTTGTAGTTT
33301 GCAACCTTGT GAATGTATAA GGACCACTGA ATTGTCCACT TCAAAAGGGT
33351 GACTTTTATG TTATGTGCAT TATATCTAAA AAAAAAATCA TAATTAGGAA
33401 GCAAGATTGA CTTCTAAGAA AAAGCGGAGT GAAATTGTTG TTTTGTGGTG
33451 AATAAATTGG GTGGGTGGGT CGCAAGAGTT TTGCTGATTA GTGATTAGAA
33501 AAATTATTCA TAATCATTGA AAATATAAAA TATTTTCTA TATGATGTAT
33551 GTAAAGAATT TGGCAAGAGA TGATGTTTGG AAAAAATAAA GAATGGCTAT
33601 TGTAGAGATC TTAAGGAAAG AAACACAGT TAAGTAGTGC TTTGTAATCA
33651 GAATATGAAG TAAGTACTGA AAGTGGATGG AGTGGCTGTT GTCAGCATGT
33701 TATACTTTAT ACATTTTATT CATAAATTTG GACTGTAGAT AAAAGTAAAC
33751 TTTTTTTTTT TTTACTCTTG AACAACAGTT TTTTTTTTTT CACTTAGACT
33801 TGCATCTGCT CCACTGAACA ATACATTTAA TTGTTAATTA TTTCCCTTT
33851 CAGGATGATA CATATACAGA AAGCTACATC AGCACAATTG GTGTGGATTT
33901 CAAAATAAGA ACTATAGAGT TAGACGGGAA AACAATCAAG CTTCAAATAG
33951 TAAGTGACTT GGCTAGTAAT TTTTTTGAAA TTTATTTTGG TAAATTTGTA
34001 ATGTATTGTT ATTTTGTATA TATTTACTAT GCTAACAAAA TTGAATGTAA
34051 AATGTCTTAA GATTATGTA CTTAAGATAG AATGGTAGAA TAAGAATTAC
34101 TTAGATTAAA AATAATATTT TCAAGATTAC TTAAGCCTCA TTGAATTTTC
34151 TGTTTCATGAA GCAGAGAAAC TCATGTTTTA AGTCAAACCT GGCTCCTATC
34201 TTTTCTTTT ATCAGTGGAA ATCTAAGTTC AAGTTTACCT TGTCTTACAC
34251 TGCAAATGTT ATAGACCATT TTTGTTTGT TTTTACTGTG CTAAGTGCAT
34301 GGAACATTAA AGGAACCTTA GGAAGAGATT CTTATATGTT GGCTCAGTTG
34351 AAGAGAAGTA CTTATGTAGT TCTAAGTATT TTTATTAGAT AGTGTGCACC
34401 AACTCTGTAG AAACACAGAA TTTTGTGGGA AAAAGGAACT TAGTTTTTGT
34451 AACATGTTCA TTTTACTGCT CAAAAAAGC AATGCTGAAA GATTTAATGA
34501 CTTGCCATACA GTTACTGGTA GAACCAAGTG ACCGAAGCTC TGTCTTCAAT
34551 ATTTTGTGTC TGTGTGCCAT CCTATCCCC TTATCCATCT TTACACCCCC
34601 AGCCCCCAAT TAAATATAGG CAATTATAAT AGTTCAGTTG TGCCTCTTCA

FIGURE 3, page 11 of 21

34651 34701 34751 34801 34851 34901 34951 35001 35051 35101 35151 35201 35251 35301 35351 35401 35451 35501 35551 35601 35651 35701 35751 35801 35851 35901 35951 36001 36051 36101 36151 36201 36251 36301 36351 36401 36451 36501 36551 36601 36651 36701 36751 36801 36851 36901 36951 37001 37051 37101 37151 37201 37251 37301 37351 37401 37451 37501 37551 37601 37651 37701 37751

GTATGGGTCT GAGTCCTGTC AGTGTGGGCA TATCTGTGGT CTTTTAAAA
ATAAATCTCT CAGTATTTTT CAGAGTAGGC TATTAGCAAG AAGTAGGCTA
TAAACACAGG AAACCAGTGA CTGCCCTTT TCATGGAACAT GATGACACAT
GGAATTGGAA GGAGTCCTGC ATTAGGAGTC AGAAGACTTA GATTTGTGT
CTTGGTTCTA GTATTTACCT GTTAGAGAA CATGGGTTTG TGTCTCTGGG
GAAAAGGCCG AAGTAACCCCT GAGACCCAGT TTCCTTTCTA AAATGTGTGT
GATGACACCT GATTTACTAA TTTATAAGCT AGTTGTGAGA ACCAACTGTA
ATAGCTTTGT GTATGTGACA ATACGTGTGA AAGCCCTTTG TAAACTTTTG
GGCAGCATAT AGATACTACT TATGATATGA CATGCCCAGA TAAATGGGTG
TTTGATAGGT TAAGTTGCTC CCTTTTCTTA CATGACTCTG ATGAGGAAAA
GAAGGTATGT TAACAAAAGA TAGGTGGCTG TGGATATTGA TATAAGTAAA
CACACTTGAT GTGTCAAAT AGGACTTGCA AGGATTTAGT TTTCAGAAAT
AGCTTGAAAT ACTTTCAATC AGTGAACAAA TTACCCTCCA TATTTTTTCC
35301 CACGATATAA GTACAGTCTC AACCTTTTAT TTGGCACCAT AAAGAGCACA
35351 TAAAGATCTA CCCAAACTG TACTTTAAAG CACTGGTATG GAATAATGT
35401 ATTATGTGTG ATCATTGGTG TTTATAAGAT TTGGGTGTGT ATTCTGTGT
35451 GAAACATPCA TATTTTGTTA CTTTCCTGTG GCTGGAAGGG ATCTTATAGG
35501 ACACTGTCTT TCATCTTTGT CTGTCTTTCA TCTTAAATAG GAATTTCTTT
35551 TCCATGCCCTG AAGGCCTCAT TTTGAACATT TTGTTTGTGT GTTTTTTAT
35601 TTTTTGAGAT ACAGTATTGC TCTGTCTCCC AGGCTGGAGT GCAGTGGCGC
35651 GATTTGAGCT CACTGCAACC TCCGCCTCCT GGGTCAAGT GATTCTCCTG
35701 CCTCAGCCTC CCTAATAGCT GGGATTACAT GTGTGTACCA CCATGCCCGG
35751 ACAATTTTTT TTTTTTTGAG ATGGAGCCTT GCTTTGTGCG CAGGCTGGA
35801 GTGCCAGTGG TGCAATCTTG GCTCGCTGCA GCCTCCGCCT CCCAGGTTCA
35851 AGCAGTTCTC TTGCCTCAGC CTCTGAGTA GCTGGGATTA CAGGCGTGCG
35901 CCACCACACC CTGCTAATTT TTTGTATTTT TAGTAGAGAC AGAGTTTCAC
35951 CATGTTGGTT AGGCTGGTCT CGAACTCCTG ACCTCGTGAT CTGCCTGACT
36001 CGGCTTCCCA AAGTGCTGGG ATTACAGGCA TGAGCCACTG TGCCAGCCT
36051 TCCGATAATT TTTGTATTTT TCGTAGAGAT GGGATTTGCG CATGTTGGCC
36101 AGGCTGGTCT CAAACTCCTT ACCTCAAGTG ATCCACCCGT CTGGCCTCC
36151 CAAAGTGCTG GGATTACAGG CGTGAGCCAC CACGCCTGGG TTTTGAACA
36201 TTTTTAAGAA GCTTACCATT TTTTCGAAAT AGCTAGTTCC ATTTACACA
36251 TAACCTCAGC TAGGCATGTT GCCTCATGCC TGTAAATCCA GCACTTTGGG
36301 AGGCCGAGGT CAGAGAGTCA CTTGAGGCCA GGAGTCAACA TAGCTCCTGT
36351 GACCAGCCTG GTCACATAG AGACTCTATC TCTACCAAAA AAAAAAAAAA
36401 AAAAAGTAAC CAGGTGTGGT GGTCATGCC TGTAGTCTTA GCTCCCAGG
36451 AGACTGAGGT GGGAGGAATG TTTGAGCCCA GGACTTCAAG GCTGCAGTGA
36501 GGCAAGATTG CACCATTGCA CCCCAGCTTT GGGGACAGAG TGAGAGACCC
36551 TGCTCAAAA ACAAAATAAG GCTGGGCGCA GTGGCTGTCC GGGCGTCTGT
36601 GTTCACGCTT ATAGTCCTAG CACTTTGGGA GGCCAAGGTG GGCAGATTGC
36651 CTGAGCTCAG GAGGTCTAAG ACCAGCCTGA GCAACATGGC GAAACCTCAT
36701 CTTTGCAAAA CATACAGAAA AAAACAAAAA AAACCACAAA ACCTCTAGTT
36751 GCCAGTTATT TTTTTTATTT ATTCTAGTG ATCTTCTTT TTTCTTTTT
36801 TCTGAGACAA AAATTTCACT TTGTCTCCCT CGCTAGAGTG CAGCGGTCAG
36851 CTCACTACAT GATTCTTTTA GAGACATGTT AATTCTTTAT ATTGAGCTGA
36901 AGCCTGTTTC TTTTACTTCT GTCTCTTCTT ATCTCTCCGC CTGTAGAGC
36951 TGCCTGAATC AGATTAATTC CTCTTTTATT GGCAAGCCTG CCCTTCAGAT
37001 TGATCTTATC ACAACCTTTC TTCTACCTCT GAAGTCTCA TTCTTTCCTG
37051 TAATGATATT TTCAGAACCT TGTGCAATTT GGGTTATTCT TACATTTTAT
37101 AAATGCCTTT TATTAAATTT GATTCTTAA ATCAAGTATG AGATATAACA
37151 CATGAGGTAA ATCCTGTCTT GATTTGGAGC CTGAATGAAT TTCTCTCTTG
37201 AACTTCAAGG GCTCATGGCC CTTTCTTATT ATTAATCAAA GACAACCATT
37251 TGTTGTTTCA GTAGCTATAT TATTCTAGT TTGGGTCTTA AGGTTTTTGA
37301 TTTGCTTGTT TTTTCTTTTT TCTTTTTTTT TTTTGTGAGA CGGAGTTTCG
37351 CTCTTGTTGC CCAGACTGGG AGTGCAATGG CGTGATCTCG GCTCACTGCA
37401 ACCTCCGCCT CCCAGGTTCA AGCGATTCTT CTGCCTCAGC CTCCCTAGTA
37451 GCAGGGATTA CAGGCATGTG CCACCACGCC GGGCTAATTT TGTATTTTAA
37501 GTAGAGATGG GGTTCCTCCA TGTGGGTAC GCTGGTCTCG AACTCCCAGC
37551 CTCAGGTGAT CCGCCTGCCT TGGCCTCCCA AAGTGCTGGG ATTACAGTCG
37601 TGAGCCACGG CGCCTGGCCG ATTTGCTTGT TTTTAATTAA AATAGGGGCC
37651 TTGGCCAGGT GCAGTTGTTT ACCCCTGTAA TCCCAGTACT TTGGGAGGCT
37701 GAGGCAGGCA GATCTCTTGA GTTCAGGAGT TCAAGACCAG TATGGGCAAC
37751 ATGGTGAAAC CCTGTCTCTA CAAAAACAC AAAATTCAGC CAGGCATGGT

FIGURE 3, page 12 of 21

37801 GGTGTGTCCC TGTAGTTCAA GGTACTCAGG AGGCTGAGGT GGGAGGATTG
 37851 CTTGAGCCCG GAGATGGAGG TTGCGGTGAG CCAAGATTGT GCCATTTGCA
 37901 CTCTAGCCTG GGCAACAGAG CGAGACCTTG TTTCAAAAAA AAAAAAGAA
 37951 AGGGTCTCAC TTTACTACTT TGTGACTGGT GTTTTAAAAA TCTAAACACA
 38001 GGCCGGGCAC GGTGGCTCAC GCCTGTAATC CCAGCACTTT GGGAGGCAGA
 38051 GGCACGCAGA TCACAAGGTC AGGAGTTCTG GACCAGCCTG GCCAGCATGG
 38101 TGAAGCCCAT CTCTACTAAA AATACAAAAA AATTAGCTGG GCATGGTGGC
 38151 AGGTGCCTGT AATCCAGCT ACTTGGGAGG CTGAGACAGG GGAATCACTT
 38201 GAACCCAGGA GCGGAGATT GCAGTGAGCC AAGATTGCGC CATTGCACTC
 38251 CAGCCTGGTG ACAGAGCGAG ACTCCGTCTG AAAAAAAAAA AAAAAATCT
 38301 AAACACAAGA TTTTACTTTT AATCCTATCA TTTCTCTTG CTGGCTTCA
 38351 GTAATCCTTC AAGTTTCTTA GGTCTTTTCA AAATCTTGAT TCTGTTGATT
 38401 TATATTTTAA TTATCTTTTC CTTTCAGCTT TTCTGTGTTA GGTGTGACAT
 38451 CTGGGTCTTT ATCTGAGTTT TATTAGATTA TAAAACATTC AGCAAGATAG
 38501 GGCAGGTAAT GAGTCCAGTT GTACACCATG GAAGGCCTCT TTCTGTGATT
 38551 GTTCATTCAT GAGGCTTTAT GAAAATGTCT ACATTACACC AGGCACTTGG
 38601 AGGTTACAGA GATGAATAAA ACATAGTCCA TTAGGAGGCA GACAATGGGA
 38651 GAGACAAACA TGGGAAAAAG TTAATCTGAT TATGAGGAGT AATGAGAATT
 38701 ACATATGAAG GAAAGTATTG TTAGTACTGT TAGGATTAG TGTGAGAAA
 38751 GTTTTCAGAG TAGCAAGGAA ACATCAGAAA TTTTACTCTT TCTGCCAGGC
 38801 ATGGGTGATG TATTATCTGT TTCTCACACT GCCACAAGGA ACTGACCAAA
 38851 ACTGGGTGAT TTATTAATAA AAAGGTTTAA TTGACTCATA GTTCTGCATG
 38901 GCTGAGGAG CCTCAGGAAA CTTACTGTGG CAGAAAGGGA AGCAGGCACG
 38951 TCTTACATGG CAGGAGGCGA GAGAGTGTGA AGGAAAGTAA GGGGGAAGAG
 39001 CCCCTTATGA GACCATCAGA TCTTGTGAGA ATTCATTAC TATCACTCGA
 39051 ATGGGGGAAA CCGTCGTCAT AATCCAATCA CTCTCCATA ATCCAATCAC
 39101 TTCCCTCAGT GATTACAAC TGAATGAGA TTTGGGTGGG GACACAGAGC
 39151 CAAACCATAT CAGTGCCTGT AGTCCAGTT ACTTGGAGGC TGAGGCAGGA
 39201 GGAACACTTG AGCCAGGAG TTCAAGATCT GCCTGGGCAA CATAGCAATA
 39251 CCTCCATTTT GGATAAAAAG GAAATTTTAC TTTTGGGTG CCATTGCTTA
 39301 GTTTAATCAG CTGTAACCTC TTGTTGACTT TTAGTCAAAA AACAAATTTT
 39351 CTTCTATCTT TTGTGAAAGA GGTGGTGAG CAAGGAAGAA AAGGAACTT
 39401 GCTTTATTGA GCAGCTTCTA TAGTCAGGCA CATTTTACAA ACATTAGTTC
 39451 ATTTAAACCC CTTTAGCTGT TGTACAAGGT GAATGCTATC TAGCATTTAC
 39501 AGATGAAGAA ACTGTTAGGT GACTCTCCCT AATATTAAAT AACCAGGAAC
 39551 CTGGATTTGA TGTTTTGAAG TCAGGGTAGC TTGATCCTCG AGTTCATGCT
 39601 CCTCCAAGG ATACACTGAA AGACTTTGAG CCTCTTTTCT TTTTCTCTC
 39651 TTTTCTTGA ACAGGATCTG GCTCTCTTGC CCAGAGTGCA GTGGTGTGAT
 39701 CTCAGCTCAC TGCAACCTCT GCCTCCTGGG CTCAAGCGAT TCTGCCTCAG
 39751 CCTCTCGAGT AGCTGGGACC ACAGGCGCAC GCCAGCATA TTGGCTAATT
 39801 TTTGGATTTT TAGTAGAGAC AGGGTTTAC CATGTTGGTC AGGCTGGTCT
 39851 CGAACTCCTG AGCTCGTAAT CCGCCCGTCT CGGCCCCACA AAGTGTGGG
 39901 ATTACAGGCG TGAGCCACCG ACCAGTCCC AACAGTTTCT TAAAACCCAG
 39951 AACTATAATG CAATAATGTT AGCATTTGTT TTGGGAGTTT GAGCCTAAAT
 40001 GGTGTAAGTG CAGTAAATG TTCTTAAAT ACGTTTATG AAAGTATTTG
 40051 GAGTCTCTTC CTTACATTTT TTTCTCTAGC ATGAAGACAA CACCTAGCCA
 40101 GGCATGGTGG CTCATGCCAG TAATGCCAGC ACTTTGGGAG AATGAGTTAG
 40151 GATAATTGCT TGAGTCCAGG AATTGAGAC CAGCCTGGGC AATGTAGCGA
 40201 GACTCTGTCT CTACAAAAA GAAAAATTA GCCGGGTGTG GTGGCATGTG
 40251 CCTGTAGTCC CAGCTACTCA GGAGGCTCAG GTGGAAGGAT TGCTTGAGGT
 40301 GGGAGGTTGA GGCTGCAGCG AGCCATGATC ATGCCACTGT ACTCAGCCTG
 40351 GATGACAGAA TGAGACGCTG CTTGAGAGGG GAAAAAAAAG ACACCTGCTT
 40401 GGGATGATTA AAGTCTGTCT TTGACTGGTA GTTATTGAA TTAGGTCCCT
 40451 CCAGTGTCTT TAATCATGGT AGAATGTGCT AGCAAGTGAG TTTGTCTTAC
 40501 ATGGAAGAGT TCTGTGTTCA AGGGCTTTCG GCCAGTGGA TTTCTTAAAC
 40551 CAGTGTAAAA GCGGTAGGG AATGTGAAAA GTATGACATA GTTCTGTCTC
 40601 TCAACAGCTT GTAATTTTAG TATTATTATC GTAAGCTCAA TTGTAGGTAC
 40651 TACTTCTTTT CTGGACTTTC AGGTGCTTAT TACCGTGCAA TTTAGTGGTA
 40701 TGAGTTGAGG ACTAATGTTT CTATATCACA TCCTGATAAT CTCCACAGTT
 40751 ATGAAACTAA AACTATTTCC CCTCCCTCCT ACACCTTTCC CCAACTTTAT
 40801 TTTAATGGAA TTGTTGGGAT TTCTTGATTG TTTTGTAAATA GTGGGACACA
 40851 GCAGGCCAGG AAAGATTTCG AACAAATCACC TCCAGTTATT ACAGAGGAGC
 40901 CCATGGCATC ATAGTTGTGT ATGATGTGAC AGATCAGGTA AGTTCCAAGA

FIGURE 3, page 13 of 21

40951 41001 41051 41101 41151 41201 41251 41301 41351 41401 41451 41501 41551 41601 41651 41701 41751 41801 41851 41901 41951 42001 42051 42101 42151 42201 42251 42301 42351 42401 42451 42501 42551 42601 42651 42701 42751 42801 42851 42901 42951 43001 43051 43101 43151 43201 43251 43301 43351 43401 43451 43501 43551 43601 43651 43701 43751 43801 43851 43901 43951 44001 44051

GGAGATTGTG TTACAGTGAC CAAGTAGGAA GCCATTATTT GATTAATGTC
AGATTCATTT ACTACTTCAT ATATAAGCCA TCAGTATTAA TTTTATGGCA
GAAACCTTTG TCCACTCTCA AATATAAATG TGAATCACCT AAAAGACATT
TGTTTTCTCTG TAATAAATAA AAGATTAGTA ATTAGTTTTA CGTTTGCTTT
CAAGGGATTG TGGTTGTATT TATTGTCAAC TAAATAACTT TGATCAAATA
GCCAAGACTC TAACATATAG GCAAGAGTTT GTAGGGAATC GTGAGTTGCT
TGGCTTATAC TGTGTTCTTG GTGTTAAGTA TTAACAGGAA TATGGCCTGG
TAATTAGAAC TTGTCCATCA GAATTGCCAA AAGTGGGATT CGGGGGTCTC
TGCCTATGGA GGATGTGGTT CAGAAATAAA GAATTTGAAT AGGATAAGCT
GTAGGAGGAT CTTAGTATGA GAATGAGTAT CTGAAGATTA GCTGTGAGAG
AGGGCAGAGC GATGGAGGGA ACAATGTGGG ACAGTGTGAA GCATGTGATC
CAGGGGCCAT AACTTTTTTT GTTACTATTT TTTTAAATCA GAAACTTAGA
TTTCAGTGTC CTTTCTATCA AAGAAAAGGA CAAAAGATAA ACGTTCAAAA
TTGGAATTTA TTTTCTTTT GGCAAATGTT AAATCTCACC TCTAATGAGA
AATCATAGCT AATTAGGAGA TAACCTACAT GTAAGCATTT AGATTCAGTG
CCATTAGAAG TGCTGGGTGG GTGATATCTG CAGGAGAAAA AAATGATGCT
AGTTTTAAAA ATCTCTACTA TTACCGTGAA ATATTTTTTAA ATGAAAACCT
TCGTCCTCTA AATATGACTG TGGAAAAGAA AATGAGTATA TTTAATAACA
TCTCTAGTAG TAACAGTAGG TCATCTTATT CATAAACCAA
AATTTTACCA AATTTCAGGC CAGGCGCAGT GGCTCATGCC TGTAATCCCA
GAACTTTGGG AGGCCGAGGC GGGCGGATCA CCTGAGGTCA GGAGTTAGAG
CTAGCCTCG CCAACATGGC AAAATCCCAT CTCTAGTAAA AATACAAAAA
TAGCCAGGC GTGGGGGCCC GTGCCTGTAA TCCTAGCCAC TTGGGAGGCT
GAGACAGGAG AATCGCTTGA ACCCAGCGGG CAGAGGTTGC AGTGAGCCGA
GATCGGCCCA TTGCACTCCA GCCTGGATGA CAGAACAAGA CTTTGTCTCA
AAAAAAAAAA AAAAAAAAAA AAAAAAATTA ATCAAATTTT AAAACCAGGT
TTTGTAGTAC ATTTAAATTG CATATCCCAA AGCAGTTGGG TTTGCCTGCG
TTGCAGTTTA ATATTAAGCT ATACTTCCCT TTCAAATAAG GTATTTTCAT
CGTTAAGCCT GTAAATTCTA GTTTGTCAAT GTTTAGATAT TTATAGTCAT
TTTAATATAT CTGTTTACGG CCAGCTGCAA TGGCTAACAC CTGTAAACTC
AGCACTTTTT GAGGCCAAGG TGGGCCGATT GAGCTCAGGA GTTCGAGACC
AGCCTGGGCA ACATAGTGAA ACTCCATCTA TACAAAAAAT CCAAAAAAAA
AAAGACAGGT GTGGTGGCAT GTGCCTGTAG TCCCAGCTAT CCCGAGGCG
GAGCGGGGAG GATGGCTTGA GCTGGGAGG TCGAGGTGTC AGTGAGCTGT
GATTGTGCCA CTGCACTCCG GCCTAGGTGA CAGAGCAAGA CCCTGTCTCA
AAAAAAAAAA TCTCTTCACT CCTAGCAGT GGTATTTTG TAGCTAGAGT
TGTCTCACTA GCTCTTTGTT ATTTGTCTGT TAGGTCAGGA ACGATGTTTC
TGTTTATTCC AGAACTATAT TATCGAACTA TATTATCAGT CTTTCAAATG
TCTTTTATAG AGTCCTTCAA TAATGTTAAA CAGTGGCTGC AGGAAATAGA
TCGTTATGCC AGTGAAAATG TCAACAAATT GTTGGTAGGG AACAAATGTG
ATCTGACCAC AAAGAAAAGTA GTAGACTACA CAACAGCGAA GGTATGTTTA
AAGTTTAATT TTCATACTGA ATTTGAAGGT GTTGAATTAT GTATGGGTTC
TGCAGTAACA GTAAGGCCAC AGCCTTTTAA AAATATGTGC ACTAGAATAC
TGTGACAGTG ACAATTTGTG TAGCATCTGT TTGGATCCAA TGAACCTAGT
TCTTACGCTT CCATTATGGA TGGTAGAAAT GCAGTAAGAA TTAGTGAAAA
AGATTTTTC A GTGTTAATTG TGCCTCATTA TTCTCTTAGG AATTTGCTGA
TTCCCTTGG A ATTCCGTTTT TGGAAACCAG TGCTAAGAAT GCAACGAATG
TAGAACAGTC TTTTCATGAC ATGGCAGCTG AGATTAATAA GCGAATGGGT
CCCGGAGCAA CAGCTGGTGG TGCTGAGAAG TCCAATGTTA AAATTCAGAG
CACTCCAGTC AAGCAGTCAG GTGGAGGTTG CTGCTAAAAA TTGCCTCCAT
CCTTTTCTCA CAGCAATGAA TTTGCAATCT GAACCCAAAGT GAAAAACAA
AATTGCTGTA ATTTACTACT ATGTAGCTGC ACTACAACAG ATTCTTACCG
TCTCCACAAA GGTGAGAGAT TGTAATGGT CAATACTGAC TTTTTTTTTA
TTCCCTTGAC TCAAGACAGC TAACCTCATT TTCAGAACTG TTTTAAACCT
TTGTGTGCTG GTTTATAAAA TAATGTGTGT AATCCTTGTT GCTTTCCTGA
TACCAGACTG TTTCCCGTGG TTGGTTAGAA TATATTTTGT TTTGATGTTT
ATATTGGCAT GTTTAGATGT CAGGTTTAGT CTCTGAAGA TGAAGTTCAG
CCATTTTGTA TCAAACAGCA CAAGCAGTGT CTGTCACTTT CCATGCATAA
AGTTTGTAGTA GATGTTATAT GTAAGATCTG ATTTGCTAGT TCTTCCTTGT
AGAGTTATAA ATGGAAAGAT TACACTATCT GATTAATAGT TTCTTCATAC
TCTGCATATA ATTTGTGGCT GCAGAAATAT GTAATTGTGT GCACACTATG
TAACAAAACA ACTGAAGATA TGTTAATAA ATATTGTACT TATTGGAAGT
AATATCAAAC TGTATGGTGA TAAGTATTGT TTTGATTCTT ATGGTTAAAG

FIGURE 3, page 14 of 21

44101 GGAAATAGAG CCTTGCATTA TATTCAACAC AGCCATTTGT GTGTGCACAA
44151 TGCAAACATAA GGTATTCTAG ACCTATCTTA GAGCAGCATC CAGTATTTGC
44201 TTTCTAGATA ATATGCCCAA TAACATGACC TAGAGGGGCT TCTGTGCTGT
44251 GTAGGGATTT AACCAACTTC AGTGGTTCAG GGAGCTCAAA CTATATGTAA
44301 AACCAAGTTTA GAATGTATGC TATCTAGCCC GTTATCTCTG ATCCTTCTCT
44351 AAAACCATTT GAAATAGCTT CATTGATCAA CATTTCATAA ATGCATCTGT
44401 GGTAGAGGTA GAAAGCAGCA CCTTTCCTAA TTGGCAAATG ATCAGACTAA
44451 TGTGTGCTAA TGTTTTTCTT CCATGCTTTC AGTCAGATTC AACTATTTTA
44501 TCCTCCACAG TTGCTTAACT TGGTGTGGA GGAGGGTTTA AGCATTAAAG
44551 TAGGAAGCAG GAAATTTGAT TGCTCTAAAT TTAGAAATTA TATCCCTAAA
44601 AATTAACAACA TGAATACTGG GTGGTAATGA TAATTGAGGC AAATGTATTT
44651 ATTTTGGTGA CATTTTGCAT ATATGAAGAT TTTCTGAAAT AGGACCTTCA
44701 AGATCCTAGG GGGTTTTGTT TGGTTTTTAA TTGTGAGGAA TAAAAAATCT
44751 TCTGCCCACA CTGGCATTTT AAGGTGACTG AGGTCAAACG TTGTTTCTCT
44801 AGGTTGAAAT AGCAGCCAAA ACATTCTTCA CGCAGGGGCT TGGGATATGG
44851 CTGTGGCAA CACATTTTGT TGTGGGCTCC TTAATTTAAT GATAAAATTT
44901 AAGTAAACA CAAGCCAAA ATGAATAGGT TTTTAAAT TTTATTTTTC
44951 ACTAAACAGG CAATTGAAAT ACATGGTACA AAAATAAGTG GTAAGATAAT
45001 TGTAATAATGA AATGGACAGA ATATCAATT TTCCATCTAT GAAAATTTC
45051 CAATAAAAAT CATAGTTTAC TTTGTATTAT AGGCGTGCTT GGTGGATCTA
45101 TTCACTCTCA CATAAGGCAA CTGACAAAT CCTGAAGTTA CCAATAGTTA
45151 TTTTGGTGAA GATCTTTAAT GCTTCAGAAG TTTTGTTTT GCCTTAATAC
45201 AGTATAAAGG GGGAAAGAGT TCAGAACTA TTTTCTAAAG TAGCTAAATG
45251 ACACAAAACA AATGTCAAGA TACTGTGATG CCATGCCGTG CACTTCATTT
45301 TTACACAGTA AAAGTTGTTT AAATTGTCAG CTTATTCTTG GTGAGTTAGC
45351 GGAAACATTA CATGAACCTA AGATGAGCAT ATTTACAGAC TTAAGTTTGG
45401 AAAATTCCAG CGTCTTTTC CCCATGGCAG TAAAGATTGG GATTTACAAC
45451 AAAATTCAGC ATGCCCTAAG ATTTGCTTCT ATGTATACGC CAATAAATGT
45501 GGTCTCGGAA AAAATATATA CCCCTTTATA CCCCATTTC CAAGTACAAA
45551 CGGTTCAAAG CTAATACAGG TTTTAATAAT CTGTCACTT AGTAAAGGGA
45601 ATTACCACTT GTTCTAAATA TAAGGTGCTG CCATAAATTA GTTTACATAG
45651 TGAGAAGAG TGTTCTTAAA TCTAAGCAGC TGCACACTCT GTGAAATCCT
45701 TTCAGAATGA TAGTCATTGT GGTCTGAGCA GTAATTTCTT ATTCTTCGAC
45751 CTTGGATTGA ATTTCCCTTA GCCTACATCT TGCCTTTCCA GCATATCTTA
45801 CCTCAAACCT TCTTTGTGTT CCATCCCAC CTAAGCTTCA AAATAGCCCT
45851 GTGTTGACGT CGTCTTCCAT TTGCTGAGCT TACCTATGGA TCTCCAAGAA
45901 CCCAGATCTT GAAACTGCTG ATCCAGCTTT GAGTATCATC ACTTCCCTGT
45951 GGATTTAACT TCCATTAATT TTAAGGGACT ACTAAGTTAT TCCAGTGTGG
46001 CATCACAGTG CAGTTAGCAA GCTCAGCTAC TTGACTCTAA TTTGGCCATG (SEQ ID NO:3)

FEATURES:

Start: 2181
Exon: 2181-2203
Intron: 2204-27090
Exon: 27091-27163
Intron: 27164-33853
Exon: 33854-33949
Intron: 33950-42859
Exon: 42860-42991
Intron: 42992-43239
Exon: 43240-43434
Stop: 43435

CHROMOSOME MAP POSITION:

Chromosome 2

ALLELIC VARIANTS (SNPs):

DNA				Protein		
Position	Major	Minor	Domain	Position	Major	Minor
397	T	-	Beyond ORF(5')			
2326	A	G	Intron			

FIGURE 3, page 15 of 21

3486	C	A	Intron
6651	-	A	Intron
8190	T	-	Intron
8281	T	C	Intron
11546	A	G	Intron
11670	C	T	Intron
11688	A	G	Intron
14938	A	C	Intron
22261	G	A	Intron
22852	G	A	Intron
27253	A	C	Intron
28098	-	A	Intron
28597	G	T	Intron
31431	C	T G	Intron
35704	C	T	Intron
35728	C	T	Intron
36690	C	T	Intron
41002	G	C	Intron
41033	A	G	Intron
43161	C	T	Intron
43765	A	G	Beyond ORF(3')
44713	G	T	Beyond ORF(3')
44831	C	T	Beyond ORF(3')

Context:

DNA

Position

397

TGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCCTCTCGGCCCACTGTAGCCTCCGCCTCCC
GGGTTCAAGCAATTTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCGCCA
CCATGCCTGGCTAATTTTGTATTTTAGTAGAGACAGTGTTCACCATGTTGGCCAGGC
TGGTCTTGAATTCCTGACCTCGTGATCTGTCGGTTTGGCCTCTCAAATTCCTGAGATTA
CAGGCATGAGCCACCGAGCCTGGCCAGTTTTCTGAGTTTTATTGAAATCAAATAAGC
[T, -]
TTTTTTTTTTTTTAATGGGCTTTAGAGTCCAGGGTAACGAACACTTTTTGGTGCCTATT
ACTGAACCATTACAGGTATTCCTGGGGTGGTGACCGTGTTTATTTCAGAAACCAACATGT
TCATTTCAGAAACCAACTCGGGTAACTTTGTATAAGTTTCACTAACTAAGGCCCATGGCA
GAATTTGAGGGCTAAGGGGTGAATTAGTGTATGGGTAGAAATAAGTGCCTTCTTTCTAT
ATTTTGGCGTTGTAGGAATTTAAAGTGATTCTGCAGTAAGTCTCAGGAGACAATTTTCTT

2326

GCTGATTGTGTTCTAGGGGACGGAGTAGGGGAAGACGTTTGCTCTCCCGGAACAGCCTAT
CTCATTCTTTCTTTTCGATTACCCGTGGCGCGGAGAGTACAGGGCGGCGGCTGCGGCAGCA
AGGGCGGCGGTGGCGGCGGCGGCGAGCTGCAGTGACATGTCCAGCATGAATCCCGAATAGT
GAGTTACAGGAGAGCACCGGTGCGCTGGGTCCGTGGGCCAGCTTGGGGGATCTTAAAGGGG
TCGAGGAGGGTTGGGGCAGAAGTCCGGGCATCGGCTGGGGTGAAGCGAGGGTGATGGGTG
[A, G]
GGAGAGGCTGGCGGCGGAGTCCGGCCCCATTGTCTGACGCGGAGGGGCGGCGCGCGG
GGGAGGGGTGGGCGCGGAGGGGTGAGCCGCGCGGCTGGACCGGGTCAAGTTAGAGGGC
CTGACTGCGGGGCGGGTGTGAGGAAGCCTGCCGAGGGGCTGGGGCGGTGTGAAGGGGT
ATCTTCTCTCGGAGGCAGTGACTTTTGAAGGAGGACTTGTCTCTAAGGGGAGGGGATGGG
GTGGGAGAGCCCTTCTAGAGGGCACTGTACAGCCCTGCGCCCGCACTCTGCGGAGCTGTC

3486

CTGGGAAGTGGTGTTCACCTCCCTTGGGTAGAGTTTGTGGGCTCTCCTCAATGGCCCTT
TAAAAATTTCTCTACAGTTTACATGCATGTAAGTAATGAATAATTGGAAGAGACCGAA
TTGGTATTCTTTTCAAGTGTCAAAGGCCCTTGAGGGATGGGGGAAAATCAGTATTTGTG
TAAAAGTTGAGTTTATTTGCTGGTTGGTCAATTACTGCTAGACATTTTCCCTAAAAGG
TCCACCCACCAAGTTTAGCTGACTGTCATATGTGTGTCACATGGCTCTGCAAAATGCTTA
[C, A]
AAGTTTTGTAATAGTGTGGCTTGAAGCTGAAATCTTTTGCCTAAACAGAAACCGTAGTA
TTTTATTAGAATTTCTAGCTTTAGAAGTTGAGGGTAGTGTCTTGTAGTGACATTTGCTG
TGTTGACAGTTTAAAAAATTTTTTTTCAAGGGCTCCAAGGACAAAGTTGGTTTTGCAC
AGTTGAACGGAGGTGAAGTTGAGGTTCTTAATTTAGTAGTTTCTTGGTAACAATAAAGA
ACATGGATTACTGCTTTATCGAGGTTTATAGACCTCTACTGTTCAAGAAATTTTCTGAA

FIGURE 3, page 16 of 21

6651 TTTCAGCACATTAAGAAATGCTTAACATGGCCAGGCGCAGTGGCTCACGCCTGTAATTCT
CAGCACTTTGGGAGGCGGAGTGGGCGGATCATTTGAGGTCATGACCAGCCTGGCCAACA
TGATGAGACACTGCCTCTACTAAAAATACAAAAATAGCTGGGTGTGGTGGTGCACGCCT
GTAATTCCAGCTACTCAGAACCTGAGGCAGGAGTCACTTGAACCTGGGAGGCGGAGG
CTGCAGTGAGTCCAGATCATGCCACTGCACTCCAGCTGAGGGACAGAGTGAGACTCCTC
[-, A]
AAAAAAAAAAAAAAAAAGAAAGAAATACTTAACATTATTCTCGTGATTATTCTCATAAC
ATTTTTTCATAATCCACTGGCTTCCAGTGGATTTTTTTAGTGTCAAGAAAATAATTTTGAT
TGGTTCATCTTTAAGGAATGTGTTAAGAATAAAGCATGTCTACCTGTCTTCAGTATACCA
GCTAACTATAGTAGGAAGAAATATAGTAGTCTACTTAGATCAACTATAATCTTTAATGC
AGAAAAAGTTTAAAGTATTACCTTATTTTTAGCCCCATCCCCTTAAGTATATCATGGC

8190 AGACCGGCTGGCCAATGTGGTGAACCCCTGCCTCTACTAAAAACACCAAATTAGCTAGG
CGTGGTGGTGTGCGCTTGTAGTCCCAAGCTACTGAGGAGGCTGAGACAAGAGAATCGCTT
GAATCTGGGAAAAAGAGGTGCGCTGAGCCAAGATTGGCCACTGCACTCCAGCCTGGGTG
ACAGAGTGAGATTCTGTCTCAAAAAATAAAAATAAAAAATTTCCCCCTTAAATCAAAT
AAGTTAAATGAGGGATGTTAGACAGTTTTTAACCATCAAATATTTTAGTTAGTTTTTT
[T, -]
TTTTTAACGTTGTCTTAAAGATGGAAGTGCTTCAAATCAAATCTTCCTTGCCAGTTCTC
TACTTGGCTTCTTTTTTTTCTTTTTGAGATAGAGTCTCACTTTGTCACTGGAGTGCCTT
GGCGTGATCTCGGCTCACTGCAACCTCCGCTTCCAGGTTAAGTGATTCTTCCACCTCA
GCCTCTCAAGTAGCTGGGAGTACAGGTGTGTGCCACCACCCCGGCTAATTTTTGTAGTT
TTAGTAGAGACAGGGTTTCACTATGTTGGCCAGGCTGGCCTCAAACCTCCTGACCTCGTGA

8281 CTGAGGAGGCTGAGACAAGAGAATCGCTTGAATCTGGGAAAAAGAGGTTGCCGTGAGCCA
AGATTGGCCACTGCACTCCAGCCTGGGTGACAGAGTGAGATTCTGTCTCAAAAAATAAA
AAATAAAAAATTTCCCCCTTAAATCAAATTAAGTTAAATGAGGGATGTTAGACAGTTTTT
AACCATCAAATATTTTAGTTAGTTTTTTTTTTTTTAACGTTGTCTTAAAGATGGAAGTGC
TTCAAAATCAAATCTTCTTGCAGTTCTCTACTTGGCTTCTTTTTTTTCTTTTGAGA
[T, C]
AGAGTCTCACTTTGTCACTGGAGTGCCTTGGCGTGATCTCGGCTCACTGCAACCTCCGCC
TTCCAGGTTTAAAGTATTCTTCCACCTCAGCCTCTCAAGTAGCTGGGAGTACAGGTGTGT
GCCACCACACCCGGCTAATTTTTGTAGTTTTTAGTAGAGACAGGGTTTCACTATGTTGGCC
AGGCTGGCCTCAAACCTCCTGACCTCGTGATCCACCCACCTCAGCCAAATGTCTGGGATTA
CTTGTGTGAGCCACGCGCTGGCTTCTACTTGGCTTTTAAAGGGAATTTTGCTTTCTGAG

11546 GTTACATTTAACCATTATATGGTGTGTAGCCATACTCACGTTACATTTGATGCATCTGC
TCCCTTTGTGTCTATATACTCATATAACATTTTGCATAAAGTTATAGGCAGTTACACCA
AGGCTGTTTCATGAACCTCAGATTAAGAATACTTGATTAGGAGATTGAAAACAGAAAAGA
GAATGTTAACTATCATTATCAATATTAATGTGAAATCTGAGAGTGACAAAGCTTAGC
TTTAAATCTGGTATCCCAAACCTATTGAGTTTTTTTTTTTTTTTTTTTTTTTGAGAC
[A, G]
AGGTGTGCTTTTGTCCCCAGGCTGGAGTGTAGTGGTGTGATCTTGGCTCACTGCAACCT
CCACCTCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCTGAAGTTGTCTGGGATTACAGG
CTGCGCCACCACGCCAGCTAATTTTTTTGTATTTATAGTAAAGACGGAGTTTACCTTAT
TGGCCAGGCTGTCTCAAACCTCCTGATCTTGTGATCCTCCGCTCGGCCTCCCAAAGTG
CTGGGATTACAGGTGTGAGCCACTGTTCCCGCCTAATTTGAGTTTTTAAATGTGGAGTT

11670 TGTTCATGAACCTCAGATTAAGAATACTTGATTAGGAGATTGAAAACAGAAAAGAGAAT
GTTAACTATCATTATCAATATTAATGTGAAAATCTGAGAGTGACAAAGCTTAGCTTTA
AATCTGGTATCCCAAACCTATTGAGTTTTTTTTTTTTTTTTTTTTTTTGAGACAAGG
TGTGCTTTTGTCCCCAGGCTGGAGTGTAGTGGTGTGATCTTGGCTCACTGCAACCTCCA
CCTCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCTGAAGTTGTCTGGGATTACAGGCTG
[C, T]
GCCACCACGCCAGCTAATTTTTTTGTATTTATAGTAAAGACGGAGTTTACCTTATTGGC
CAGGCTGGTCTCAAACCTCCTGATCTTGTGATCCTCCGCTCGGCCTCCCAAAGTGCTGG
GATTACAGGTGTGAGCCACTGTTCCCGCCTAATTTGAGTTTTTAAATGTGGAGTTTAAAG
ATGTTAGTCTTAAAGTGGGTAGATGAAATTTATAAAAAATAGTCAAATAGCTAAATTTAT
AAAAGGCCATTGAAACAATTTTGTGAAATATATAATGTGGATAATTATGTAGTGCTTTA

11688 TAAGAATACTTGATTAGGAGATTGAAAACAGAAAAGAGAATGTTAACTATCATTATCAA
TATTAATGTGAAAATCTGAGAGTGACAAAGCTTAGCTTTAAATCTGGTATCCCAAAC

FIGURE 3, page 17 of 21

14938
22261
22852
27253
28098

CATTTGAGTTTTTTTTTTTTTTTTTTTTTTTTTTTGGACAAAGGTGTCGCTTTGTCCCCAG
GCTGGAGTGTAGTGGTGTGATCTTGGCTCACTGCAACCTCCACCTCCAGGTTCAAGTGA
TTCTCCTGCCTCAGCCTCTGAAGTTGCTGGGATTACAGGCTGCGCCACCACGCCAGCTA

[A, G]

TTTTTTGTATTATAGTAAAGACGGAGTTTACCTTATTGGCCAGGCTGGTCTCAAACCTC
CTGATCTTGTGATCCTCCCGCCTCGGCCTCCCAAAGTGTGGGATTACAGGTGTGAGCCA
CTGTTCCCGGCCTAATTTGAGTTTAAAATGTGGAGTTTAAAGATGTTAGTCTTAAAGTGG
GTTAGATGAAATTTATAAAATAGTCAAATAGCTAAATTTATAAAGGCCATTGAAACA
ATTTTGTGAAATATATAATGTGGATAATTATGTAGTGCTTTATGTGTAGATTGGTGGTTA

14938

CATGGTAGTGTGCACCTGTAGTCCCAACCACTTGGGAGGCTGAGGTGGGAGGATTGCCTG
AGGCCAGGAGTTTGAGACCTGGGCAGCATATGAAGACCTGTCTCTAAAAAATAAAAAAT
AAAAAATAGCCAGGTGTGGTTGGTGTGCTTGTGGTCCAGCTACTCAAGAGGCTGAGGCA
AGAGGGTTGCTTGAGCCCAGAAGTTGGAGGCTGCCGTGAAGTGTGATTGCACCACTGCAC
TTCAGCCTGGGTGACATAGCAAGACCTGTCTCTGTGGTGGTGGTGGGTGGGGTGGGG

[A, C]

AGGGATTTAAGAAGGTTTGTGAGGTATGTATTATTTATAAATGGGCTTTAACTTTACC
CTTCACATCTTGGGTGAAATTAATTGTATCCATCTCAGTTTTTCTGTCTTGTATATA
TTTAACTTGGAGACTTAGAGGTGATGGATGTCTTCTATGAAAAGCAATGAAGCAGAG
GGCTGCCTTCTCTGTGTAGAGGACACTTGTGTCAGAGCATGTTACTGTTTATGCA
TTGCTAGGCTTTGGGAGTTGTGACTTGTATGATCATAGTACTTACAACTATTAGTTGGCA

22261

CACCCACAGATAGCTATGTCAAACGTAAGGGTGGAGAAACACAGACCCCAAACTTCTCGA
GGGTAGAAAAATATGAGGTTATAGTAGATTAGAACTACAAAAGCTAGAGGAAGTTCTGAA
CTGGAACAGTGGATAGGATTACTAGAATAATTACGAGGGTGACAATTGTAATCTTC
ATAGGTTTCTTTTTTTTCTCTTTTTTTTTTTTTTTTGTAGATGGAGTCTCGCTCTGTTG
CCCAGGCTGGAGTGAATGGCGCAGTCTCTCCTCACTGCAACCTCCGCTCCTGGGTCCA

[G, A]

GTGATTCTCCTGCCTTAGCCACCCAAGTAGCTGGGATTACAGGCATCTGCCACCATGCTG
AGCTAATTTTGTATTTTTTTTTTTTTTTTAGTAGAGACGGGGTTTACCATGTTGGTCAGGCTG
GTCTTGAACCTCCTGACCTCAGGTAATCCACCCACCTTGGCCTCCCAAAGTGTGGGATTA
CAGGTGTAGCCACCGCGCCAGCCAAATTTTATTGGTTTCTAAACTAGCGTAATTTAG
TTTTTTTCACTTAAGTCAAATATATTATTGTAGGATAAAAACTTAGTGATCCAAATTC

22852

ATCCAAATTCATGAGGAATGAAGAATAAATACATTTAAAGTCTTACCATTGCTAAATTA
GTCTTGGCTCTTTGTACCAAAATCTGTCTTGTGCTCTGTAATTTTATATTGTATATT
TTCTATCAACATTTTACTGTGTGGTGTGTTGTAATTAATAAAACGTTTAAAGCAAAC
TCAGAACAAATGAATCTCACGAATATTCAGTATATTACAGTTGAGAAATAAACTACTTC
TGTTAGTGGTAATTTAAATGTCCCAATGCAAGTTAACGTGTCACTGATCACGCTATTCA

[G, A]

GTGTGTGCTTTGATAAGGGGAGGTGGGAAGTTTGTGGGTTGATTTTATTGCTTTT
TCATGTGACTGTTGTGATGTTAGTAAACAAATGGTTTGCAGAGAGAACCAGTAGTCTTTG
CAAAGATTGTCTTATACAGAGCACTCAATCTTCATATTATTTATAATGGCTTTAATTTA
AGCCTTAAATATTAGAACTCATAAATAATTTTTTATTGTTTTTTTGTAGATGGAGTT
TCGCCCTTATTGTCCAGGCTGAAGTACAATGATGTGATCTTGACTCACTGCAACCTCCGC

27253

GCTTAAGCCATGCATGGGCTTTATAGGAGATGTAGTCTTCACAGTGAGTTGTTATTTGTA
GCTGTGTTTTTTTGTATAGCTTATAGCAATGCAGTGTGCTTTTTATTAACATCATT
TTCTTTTCTTTTGCAGTGATTATTTATCAAGTTACTTCTGATTGGCGACTCAGGGGT
TGGAAAGTCTTGCTTCTTAGGTTTGCAGTAAGTTGAAATTGAAATGTCTTACAAT
TAATGGTACAATTAATGCTATGTATGTTTTCTAGGTAGATAAAAAATAAACAGTTTATTC

[A, C]

GAATAAGTTAATCTTCCAGAATTTATATATTTAAAGACTCCAAATATACATCCCCAGTG
GTATCTTGACTGTTAAATAGAAAAATATTGTTGCTCTTAAAGAAATTCAGTGAAGTCT
GGTTATAAAGTCAGAAATGTCTAATACTTTTGGTCAGAGTCAAACAGCAGTTCCAATATAG
GCAGCAAGTTAAAGGGTAGTTGGTGGCCTGTGTTGAAAGCGACTTGATGAAAAAATC
TTTAAATTAACTTTAGTAGAATAAAAAGAAAAAGCAGAGCCAGGTGACGAGTGGATCA

28098

CTTTAAATTTAGCATGTTTCTGGCCAGGTGCGGTGGCTCACGCCTGTAATCCCAGCACT
TTGGGAGGCCGAGACGGGCGGATCACAAGGTCAAGAGATTGAGACCATCCTGGCTAACAC
GGTGAACCCCGTCTCTACTAAAAATACAAAAATCAGCTGGGTGTGGTGCCACACGCCT
GTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCAGGAGGCGGAGG
TTGCAGTGAGCTGAGATGGTGCCACTGCACTCCAGCCTGGCAACAGAGCAAGACTGTCTC

FIGURE 3, page 18 of 21

[-, A]
 AAAAAAAAAAGAAAAAAAAATAAAAAACAAATTAGCATGTTTCCCTTCTAGAGATCATTGT
 TTCTCAGAGCATGGACCAAAGACTCCTGGGGGTACCAAGACCCTCTCAGGTAGCCCATG
 AGGTCAAAATATCCTAATAATACTAAGATGTTAGTATTTGTAAGGAAATATTTACTTGGT
 AATAATACTAATAATAAAGATGTTTGCCTTTTTCAGTGATGACATTGGCTCTGGTACAAA
 AGCATGTGGGTAAAATTGCTGCTGGCTTGGTACACATCAAGGCAGCGCTAAGCTCCAAAT

28597
 GATGTTTGCCTTTTTCAGTGATGACATTGGCTCTGGTACAAAAGCATGTGGGTAAAATTG
 CTGTGGCTTGGTACACATCAAGGCAGCGCTAAGCTCCAAATTGTACTCATGGTGATGGC
 ATTCTTTACCTCTGTGCCCTCACAGGAACAAAAACAAGCCGTGCCATTTTATTGAAGAT
 TGTCTTGACAAAACAGTTAAAATGATTAATTTTGA AAAATGTTGATCCATGAGTATTC
 CTTTAAAAATATTTGTGAAGAAATGGGAAGTTCACATAAAACAATGTTTTTTTTTGT
 [G, T]
 TTTTTTTTTTTTTTTTGAGACAGATTCTGGCTGTGTTGCCAAGCTAGAGTGCAAGTGGC
 GTCTGGCTCCCAGGCTCAAGCTGTTCTCCACTTCAGCCTCCCAAGTGGCTGGGACCTCC
 CAAGTGGATGCGCCATCATGCCTGGCTGATTTTGTATTTTGTAGTGACAAGTCTC
 ACTGTGTTGCACAGGCTGGTCTCAAACCTCTGAGCTCAAGCGATGCATGTGCCTCAGCCT
 CCCAAAGTGCTGGAGAAAGCACTTTTACTGCATACTGGCTAGTGTGTTGTTATTTTGG

31431
 CTGCATTTTTTTTTTTTTTTTGGTTTGAGATGGAGTCTCGCTCTGTGCCCCAGGCTGGA
 GTGCAGTCGTGCAATCTCGGCTCACTGCAGCCTCCACCTCATGGGTTCAAGCGATTCTCC
 ATCTTGGTCTCCTGACTAGCTAGGTTTACAGGCGTGTGCCATCACACCCACTAATTTTT
 GTATTTTGTAGTAGAGACAGGTTTCCACCATGTTGGCCAGGCTGGTCTTGAACCTCCTGATC
 TAAAGTGAGCCTCCACCTTGGCCTCCCAAAGTGCTGGGATTACATATGTGAGCCACTGC
 [C, T, G]
 CCTGGCCTCTATATACTTCTATAGTACCTGATACTTATTAGGCACTCAATTACAACATAA
 CTTTTTTTTTTTTTTTTTTTTTTTGAGACAGAGACATGCCTTGTGCGCTGGGCTGGAGTGC
 AGTGGCACAGTCTCGGCTCACTGCAACCTTCACCTCCCGGTTCAAGTGATTCTCCTTCC
 TCAGCCTCCCGGTTAGCTGGGATTACAGGCGCCCGCCACCAGTCCAGCTAATTTTTTGT
 ATTTTTAATAGAGATGAGGTTTACCACCTTGGCCAGGCTGATCTCAAACCTCCTGACCTT

35704
 ATGTGTGATCATTGGTGTATAAGATTGGGTGTGATTCTGTGTGAAACATTTCATAT
 TTTGTTACTTTCTGTGGCTGGAAGGGATCTTATAGGACACTGTCTTTCATCTTTGTCTG
 TCTTTTCATCTTAAATAGGAATTTCTTTTCCATGCCTGAAGGCCTCATTTTGAACATTTTG
 TTTGTTTGTTTTTTATTTTTTGAGATACAGTATTGCTCTGTCTCCCAGGCTGGAGTGCA
 GTGGCGCGATTGAGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGTGATTCTCCTGCCT
 [C, T]
 AGCCTCCCTAATAGCTGGGATTACATGTGTGTACCACCATGCCCGACAATTTTTTTTTT
 TTTGAGATGGAGCCTTGCTTTGTGCGCCAGGCTGGAGTGCCAGTGGTGCAATCTGGGCTC
 GCTGCAGCCTCCGCCTCCAGGTTCAAGCAGTTCTCTTGCCCTCAGCCTCCTGAGTAGCTG
 GGATTACAGGCGTGCGCCACCACACCTGCTAATTTTTTGTATTTTGTAGTAGACAGAG
 TTTACCACATGTTGGTTAGGCTGGTCTCGAACTCCTGACCTCGTGATCTGCCTGACTCGGC

35728
 GATTTGGGTGTGATTCTGTGTGAAACATTTCATATTTTGTACTTTCTGTGGCTGGAA
 GGGATCTTATAGGACACTGTCTTTCATCTTTGTCTGTCTTTCATCTTTAATAGGAATTTT
 TTTTCCATGCCTGAAGGCCTCATTTTGAACATTTTGTGTTGTTTGTTTTTTATTTTTTGA
 GATACAGTATTGCTCTGTCTCCCAGGCTGGAGTGCAAGTGGCGCATTTGAGCTCACTGCA
 ACCTCCGCCTCCTGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCTAATAGCTGGGATTA
 [C, T]
 ATGTGTGTACCACCATGCCCGACAATTTTTTTTTTTTGGAGATGGAGCCTTGCTTTGTC
 GCCCAGGCTGGAGTGCCAGTGGTGCAATCTTGGCTCGCTGCAGCCTCCGCCTCCAGGTT
 CAAGCAGTTCTCTTGCTCAGCCTCCTGAGTAGCTGGGATTACAGGCGTGCGCCACCACA
 CCCTGCTAATTTTTTGTATTTTGTAGTAGACAGAGTTTACCACATGTTGGTTAGGCTGGT
 CTCGAACCTCCTGACCTCGTGATCTGCCTGACTCGGCTTCCCAAAGTGCTGGGATTACAGG

36690
 AAAAAAAAAAAAAAAGTAACCAGGTGTGGTGGTCCATGCCTGTAGTCCTAGCTCCCCAG
 GAGACTGAGGTGGGAGGAATGTTTGGAGCCAGGACTTCAAGGCTGCAGTGAGGCAAGATT
 GCACCATTGCACCCAGCTTGGGGACAGAGTGAGAGACCCTGTCTCAAAAACAAAATAA
 GGCTGGGCGCAGTGGCTGTCCGGGCGTCGTGGTTTACGCTTATAGTCTTAGCACTTTGGG
 AGGCCAAGGTGGGCAGATTGCCTGAGCTCAGGAGGTCTAAGACCAGCCTGAGCAACATGG
 [C, T]
 GAAACCTCATCTTTGCAAAACATACAGAAAAAACAAAAAACCAAAACCTCTAGTT
 GCCAGTTATTTTTTTTATTTATTCTAGTGATTCTTCTTTTTTTCTTTTCTGAGACAA

FIGURE 3, page 19 of 21

1000 900 800 700 600 500 400 300 200 100 0

AAATTTCACTTTGTCTCCCTCGCTAGAGTGCAGCGGTGAGCTCACTACATGATTCTTTTA
GAGACATGTTAATCTTTATATTGAGCTGAAGCCTGTTTCTTTACTTCTGTCTCTTCTT
ATTCTCCGCCCTTAGAGCTGCCTGAATCAGATTAATTCCTCTTTATTGGCAAGCCTG

41002 GAGTTGAGGACTAATGTTTCTATATCACATCCTGATAATCTCCACAGTTATGAAAACATA
ACTATTTCCCTCCCTCCTACACTTTTCCCAACTTTATTTAATGGAATTGTTGGATT
TCTTGATTGTTTGTAAATAGTGGGACACAGCAGGCCAGGAAAGATTTCGAACAATCACCT
CCAGTTATTACAGAGGAGCCCATGGCATCATAGTTGTGTATGATGTGACAGATCAGGTAA
GTTCCAAGAGGAGATTGTGTACAGTGACCAAGTAGGAAGCCATTATTGATTAAATGTCA
[G, C]
ATTCATTTACTACTTCATATATAAGCCATCAGTATTAATTTTATGGCAGAAAACCTTTGTC
CACTCTCAAATATAAATGTGAATCACTTAAAAGACATTGTTTCTGTAAATAAATAAAA
GATTAGTAATTAGTTTACGTTTGCTTTCAAGGATTCTGGTGTATTATTGTCAACTA
AATAACTTTGATCAAATAGCCAAGACTCTAACATATAGGCAAGAGTTTGTAGGAATCGT
GAGTTGCTTGGCTTATACTGTGTTCTTGGTGTAAAGTATTAACAGGAATATGGCCTGGTA

41033 CTGATAATCTCCACAGTTATGAAAACATAAATTTCCCTCCCTCCTACACTTTTCCCC
AACTTTATTTAATGGAATTGTTGGATTCTTGATTGTTTGTAAATAGTGGGACACAGC
AGGCCAGGAAAGATTTTGAACAATCACCTCCAGTTATTACAGAGGAGCCCATGGCATCAT
AGTTGTGTATGATGTGACAGATCAGGTAAGTTCCAAGAGGAGATTGTGTACAGTGACCA
AGTAGGAAGCCATTATTGATTAAATGTGAGATTCACTTACTACTTCATATATAAGCCATC
[A, G]
GTATTAATTTTATGGCAGAAAACCTTTGTCCACTCTCAAATATAAATGTGAATCACTTAAA
AGACATTTGTTTCTGTAAATAAAGATTAGTAATTAGTTTACGTTTGCTTTCAA
GGGATTCTGGTTGTATTATTGTCAACTAAATAAATTTGATCAAATAGCCAAGACTCTAA
CATATAGGCAAGAGTTTGTAGGAATCGTGAGTTGCTTGGCTTATACTGTGTCTTGGTG
TTAAGTATTAACAGGAATATGGCCTGGTAATTAGAAGTTGTCCATCAGAATTGCCAAAAG

43161 AGTCCTTCAATAATGTTAAACAGTGGCTGCAGGAAATAGATCGTTATGCCAGTGAAAATG
TCAACAAATGTTTGGTAGGGAACAAATGTGATCTGACCACAAAGAAAGTAGTAGACTACA
CAACAGCGAAGGTATGTTTAAAGTTTAAATTTTCACTACTGAATTTGAAGGTGTTGAATTAT
GTATGGGTTCTGCAGTAACAGTAAGGCCACAGCCTTTAAAAATATGTGCACTAGAATAC
TGTGACAGTGACAATTTGTGTAGCATCTGTTGGATCCAATGAACCTTAGTTCTCAGCGT
[C, T]
CATTATGGATGGTAGAAATGCAGTAAGAATTAGTGAAAAAGATTTTTCAGTGTTAATTGT
GCCTCATTATCTCTTAGGAATTTGCTGATTCCCTTGGAATTCGGTTTTTGGAAACCAAGT
GCTAAGAATGCAACGAATGTAGAACAGTCTTTTCATGACGATGGCAGCTGAGATTAAAAAG
CGAATGGGTCCTCGGAGCAACAGCTGGTGGTGTGAGAAGTCCAATGTAAAATTCAGAGC
ACTCCAGTCAAGCAGTCAGTGGAGGTGTCTGTAAAATTTGCCTCCATCCTTTTCTCAC

43765 AATGAATTTGCAATCTGAACCAAGTGAAAAACAAAATTCCTGAATTGTACTGTATGT
AGCTGCACTACAACAGATTCTTACCGTCTCCACAAAGGTGAGAGATTGTAAATGGTCAAT
ACTGACTTTTTTTTTTATTCCTTGACTCAAGACAGCTAACTTCATTTTCAGAACTGTTTT
AAACCTTTGTGTGCTGGTTTATAAATAATGTGTGAATCCTTGTGCTTCTCGATACC
AGACTGTTTCCCGTGGTGGTTAGAATATATTTGTTTGTATGTTATATTGGCATGTTT
[A, G]
GATGTCAAGTTTAGTCTTCTGAAGATGAAGTTCAAGCCATTTTGTATCAAACAGCACAAAGC
AGTGTCTGTCACTTTCCATGCATAAAGTTTAGTGAGATGTTATATGAAGATCTGATTTG
CTAGTTCTTCTGTAGAGTTATAAATGGAAAGATTACACTATCTGATTAATAGTTTCTT
CATACTCTGCATATAATTTGTGGCTGCAGAAATATTGAATTTGTTGCACACTATGTAACA
AAACAACTGAAGATATGTTTATAAATATTGTACTTATTGGAAGTAATATCAAACGTGAT

44713 AAGCAGCACCTTTCTTAATTTGGCAAATGATCAGACTAATGTGTGCTAATGTTTTCTTCC
ATGCTTTCAGTCAGATTCAACTATTTTATCCTCCACAGTTGCTTAACTTGGTGGTGGAGG
AGGGTTTAAAGCATTAAAGATAGGAAGCAGGAAATTTGATTGCTCTAAATTTAGAAATATA
TCCCTAAAAATTAACATGAATACTGGGTGGTAATGATAATTGAGGCAAAATGATTTAT
TTTGGTGACATTTTGCATATATGAAGATTTTCTGAAATAGGACCTTCAAGATCCTAGGGG
[G, T]
TTTTGTTGGTTTTTAATTTGTGAGGAATAAAAAATCTTCTGCCCACACTGGCATTTTAAG
GTGACTGAGGTCAAACGTTGTTTCTTAGGTTGAAATAGCAGCCAAAACATCTTCACGC
AGGGGCTTGGGATATGGCTGCTGGCAACACATTTTGTGTTGGGCTCCTAATTTAATGAT
AAAATTTAAGCTAAACACAAGCCAAAATGAATAGGTTTTTTAATTTTATTTTCACT
AAACAGGCAATTGAAATACATGGTACAAAAATAAGTGTAAGATAATTGTAAATGAAAT

FIGURE 3, page 20 of 21

44831 GGAGGGTTTAAGCATTAAAGATAGGAAGCAGGAAATTTGATTGCTCTAAATTTAGAAATTA
TATCCCTAAAAATTAAAACATGAATACTGGGTGGTAATGATAATGAGGCAAATGTATTT
ATTTTGGTGACATTTTGCATATATGAAGATTTTCTGAAATAGGACCTTCAAGATCCTAGG
GGGTTTGTGTTGGTTTAAATTGTGAGGAATAAAAAATCTTCTGCCACACTGGCATT
AAGGTGACTGAGGTCAAACGTTGTTTCCTTAGGTTGAAATAGCAGCCAAAACATTCTTCA
[C, T]
GCAGGGGCTTGGGATATGGCTGCTGGCAACACATTTGTTGTGGGCTCCTTAATTTAATG
ATAAAATTTAAGCTAAACACAAGCCAAAATGAATAGGTTTTTTAATTTTATTTTCA
CTAAACAGGCAATTGAAATACATGGTACAAAAATAAGTGGTAAGATAATTGTAAATGAA
ATGGACAGAAATATTCAATTTCCATCTATGAAAATTTACAATAAAAATCATAGTTTACT
TTGTATTATAGGCGTGCTTGGTGGATCTATTCATCCTCACATAAGGCAACTGACAAATTC

44831
44832
44833
44834
44835
44836
44837
44838
44839
44840
44841
44842
44843
44844
44845
44846
44847
44848
44849
44850
44851
44852
44853
44854
44855
44856
44857
44858
44859
44860
44861
44862
44863
44864
44865
44866
44867
44868
44869
44870
44871
44872
44873
44874
44875
44876
44877
44878
44879
44880
44881
44882
44883
44884
44885
44886
44887
44888
44889
44890
44891
44892
44893
44894
44895
44896
44897
44898
44899
44900
44901
44902
44903
44904
44905
44906
44907
44908
44909
44910
44911
44912
44913
44914
44915
44916
44917
44918
44919
44920
44921
44922
44923
44924
44925
44926
44927
44928
44929
44930
44931
44932
44933
44934
44935
44936
44937
44938
44939
44940
44941
44942
44943
44944
44945
44946
44947
44948
44949
44950
44951
44952
44953
44954
44955
44956
44957
44958
44959
44960
44961
44962
44963
44964
44965
44966
44967
44968
44969
44970
44971
44972
44973
44974
44975
44976
44977
44978
44979
44980
44981
44982
44983
44984
44985
44986
44987
44988
44989
44990
44991
44992
44993
44994
44995
44996
44997
44998
44999
45000